UNIVERSITY
OF MICHIGAN

APR 30 1952

MATHEMATICS
LIBRARY

Display
D

AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

R. BAER
UNIVERSITY OF ILLINOIS

D. C. LEWIS, JR.
THE JOHNS HOPKINS UNIVERSITY

S. EILENBERG
COLUMBIA UNIVERSITY

A. WINTNER
THE JOHNS HOPKINS UNIVERSITY

WITH THE COÖPERATION OF

L. V. AHLFORS
S. S. CHERN
W. L. CHOW
P. R. HALMOS
P. HARTMAN
G. P. HOCHSCHILD
I. KAPLANSKY
W. S. MASSEY
H. SAMELSON
R. M. THRALL
A. D. WALLACE
A. WEIL

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

Volume LXXIV, Number 2
APRIL, 1952

THE JOHNS HOPKINS PRESS
BALTIMORE 18, MARYLAND
U. S. A.

CONTENTS

	PAGE
An existence theorem of calculus of variations for integrals on parametric surfaces. By LAMBERTO CESARI,	265
The "independent scalars" in homogeneous turbulence. By N. COBURN,	296
Homogeneous Dirichlet problem for inhomogeneous ultrahyperbolic equation. By O. G. OWENS,	307
Oscillation theorems. By CHOY-TAK TAAM,	317
On the equation $a^x - b^y = 1$. By WM. J. LEVEQUE,	325
Sum functions in elementary p-adic analysis. By GORDON OVERHOLTZER,	332
A class of partially ordered Abelian groups related to Ky Fan's characterizing subgroups. By A. H. CLIFFORD,	347
A note on Lie algebras of characteristic p. By N. JACOBSON,	357
On the logarithms of bounded matrices. By AUREL WINTNER,	360
On parallel surfaces. By AUREL WINTNER,	365
The spectra of quantum-mechanical operators. By C. R. PUTNAM,	377
On non-oscillatory linear differential equations of second order. By PHILIP HARTMAN,	389
On the Dirichlet series associated with Ramanujan's τ-function. By GEORGE SHAPIRO,	401
Analytic functions of bounded type. By A. FRANCK,	410
Integral geometry in Hermitian spaces. By L. A. SANTALÓ,	423
On the parabolic generators of the principal congruence subgroups of the modular group. By EMIL GROSSWALD,	435
On solutions of linear partial differential equations of mixed type. By STEFAN BERGMAN,	444
On the applicability of linear differential operators of infinite order to functions of class $L^2(B)$. By PHILIP DAVIS,	475
Uniqueness theorems for two free boundary problems. By JAMES B. SERRIN, JR.,	492
Extensions of difference fields. By RICHARD M. COHN,	507

The AMERICAN JOURNAL OF MATHEMATICS will appear four times yearly.

The subscription price of the JOURNAL for the current volume is $7.50 (foreign postage 50 cents); single numbers $2.00.

A few complete sets of the JOURNAL remain on sale.

Papers intended for publication in the JOURNAL may be sent to any of the Editors.

Editorial communications should be sent to Professor AUREL WINTNER at The Johns Hopkins University.

Subscriptions to the JOURNAL and all business communications should be sent to THE JOHNS HOPKINS PRESS, BALTIMORE 18, MARYLAND, U. S. A.

Entered as second-class matter at the Baltimore, Maryland, Postoffice, acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, Authorized on July 3, 1918.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

AN EXISTENCE THEOREM OF CALCULUS OF VARIATIONS FOR INTEGRALS ON PARAMETRIC SURFACES.*

By LAMBERTO CESARI.

In recent years the concept of integral $\mathfrak{J}(S) = \int\int_S F(x, p)\,dudv$, $x = (x^1, x^2, x^3)$, $p = (p^1, p^2, p^3)$, on a continuous surface S in parametric form, $S: x = x(u, v)$, $(u, v) \in Q \equiv (0, 1, 0, 1)$, has been studied by the author in connection with the theory for Lebesgue area [2, 3, 4]. In particular, various necessary conditions and also sufficient conditions for lower semicontinuity [3, 4] have been obtained, which extend, to the large class of all surfaces in parametric form and of finite Lebesgue area, the well-known theorems of Tonelli [26, vol. 1] for curves of finite length and of McShane [14] for surfaces under particular conditions.

In the present paper the problem of the existence of the absolute minimum of integrals $\mathfrak{J}(S)$ for surfaces in parametric form and of finite Lebesgue area is studied with the aim of advancing for such surfaces an existence theory based only upon geometrical considerations and the concept of lower semicontinuity. The following existence theorem is obtained (no. 25):

Each positive definite semiregular integral $\mathfrak{J}(S) = \int\int_S F(x, p)\,dudv$ *has an absolute minimum in the class of all surfaces of finite Lebesgue area which are contained in a given closed bounded convex part A of the space E and whose boundary is a given closed Jordan curve C, provided that C is the boundary curve of at least one surface of finite Lebesgue area in A.*

This theorem, in which the condition of the boundedness of A can be removed (no. 42), extends an analogous theorem of Tonelli for integrals on a curve [26, vol. 2], as well as a theorem of McShane [17] for integrals on a surface $\mathfrak{J}(S) = \int\int_S F(p)\,dudv$ in which the function F does not depend upon the point x in the space E_3. Another particular case is the problem of Plateau for which Rado, McShane, Courant, Douglas, Tonelli gave, a long time ago, final existence theorems, and for which the integral involved is the area integral $\mathfrak{J}(S) = \int\int_S |\,p\,|\,dudv$.

* Received March 14, 1951.

The proof of the above existence theorem (§ 3) is based only on the recalled conditions for lower semicontinuity [3], on some results of Lebesgue area theory for surfaces, on previous observations of McShane [18] and C. B. Morrey [19], and, above all, on a detailed study of a new suitable smoothing process for polyhedral surfaces (§ 2). The author is indebted to Professor L. M. Graves for the very short and elegant proof (§ 2, n. 13) of Lemma 15, which supplants a previous somewhat longer one.

In order to point out the elementary and direct character of the procedure used in the present paper the direct application of any existence theorems for the Dirichlet problem (no. 43) has been avoided.

§ 1. Generalities on Surfaces and Integrals.

1. Let $a = (a^1, a^2, \cdots, a^n)$ be any point (vector) of the space E_n, and let $|\, a \,|$ be the norm of a; that is, $|\, a \,| = [(a^1)^2 + \cdots + (a^n)^2]^{1/2}$. Let $\{a_1, a_2\} = |\, a_1 - a_2 \,|$ be the euclidean distance between two vectors a_1, a_2; hence $\{a, 0\} = |\, a \,|$, where $0 = (0, 0, \cdots, 0)$. For any set $A \subset E_n$ let diam A be the diameter of A; that is, diam $A = \operatorname{Sup} |\, x - y \,|$ for any $x, y \in A$. For any two sets $A, B \subset E_n$ let $\{A, B\}$ be the distance between A and B; that is, $\{A, B\} = \operatorname{Inf} |\, x - y \,|$ for any $x \in A$, $y \in B$. For any set A let A_0, A^*, $\bar{A} = A + A^*$ be, respectively, the subset of the interior points, the boundary set, the closure set of A.

We have to deal especially with the space E_3 of the points $x = (x^1, x^2, x^3)$ and the space E_2 of the points $w = (u, v)$. Let $Q \equiv [0 \leq u \leq 1, 0 \leq v \leq 1]$ be the unit square of E_2 and let us fix the counterclockwise orientation on Q^* as the positive sense.

2. Let $x(w), w \in Q$, be any single-valued vector-function continuous on Q. The equation $S: x = x(w), w \in Q$, defines a continuous oriented Fréchet surface and the equation $C: x = x(w), w \in Q^*$, defines a continuous oriented closed curve C. We say, in the following, simply that S is a surface and C is a curve. We call C the boundary curve $C = \partial S$ of the surface S.

Let $\| S, S' \|$, $\| C, C' \|$ be the Fréchet distance between two surfaces S, S', or two curves C, C'. We say that two surfaces S, S', or two curves C, C', are Fréchet-equivalent, and we indicate this fact by $S \sim S'$, or $C \sim C'$, if and only if $\| S, S' \| = 0$, or $\| C, C' \| = 0$. The following statements are well known.

LEMMA 1. i) $\| S, S' \| = \| S', S \|$; ii) $\| S, S' \| \geq 0$; iii) $\| S, S'' \| \leq \| S, S' \| + \| S', S'' \|$; iv) *if* $S \sim S_1$, $S' \sim S_1'$, *then* $\| S, S' \| = \| S_1, S_1' \|$. *The same for curves.*

LEMMA 2. [*See* 6] $\| \partial S, \partial S' \| \leq \| S, S' \|$. *Hence if* $S \sim S'$, *then also* $\partial S \sim \partial S'$.

If $S \sim S'$, $S: x = x(w), w \in Q$, $S': x = x'(w), w \in Q$, we say also, for simplicity's sake, that $x = x(w)$ and $x = x'(w)$ are different representations of the same Fréchet surface.

3. [For this no. see 1 and 6] Let $S: x = x(w), w \in Q$, be a given surface. For each point $w \in Q$ we call the point $x(w)$ the image of w on S. We indicate by $[S]$ the set of all points $x \in E_3$ such that x is the image of at least one point $w \in Q$. We say that $[S]$ is the set of points covered by the surface. Given a set $I \subset Q$, we indicate by $x(I)$ the set of points $x \in E_3$ such that x is the image of at least one point of I. We say that $x(I)$ is the image of I on S. We know the set $[S] \subset E_3$ is bounded, closed, connected, locally connected. For each point $x \in [S]$ let us denote by $S^{-1}(x)$, or the counter-image of x, the set of all points $w \in Q$ whose image is x. We know that $S^{-1}(x)$ is a closed subset of Q, hence its components γ are subcontinua of Q (possibly single points of Q). Let us call G the collection of all continua $\gamma \subset Q$ which are components of at least one set $S^{-1}(x) \subset Q$. The collection G has the following properties: i) each point $w \in Q$ belongs to one and only one continum γ of G; ii) G is the collection of the maximal continua of Q on which the vector $x(w)$ is constant; iii) the collection G is upper semicontinuous on Q.

We say that a surface S is a *base surface* [in 1 an A-surface] if α) for any continuum $\gamma \in G$ the open set $E_2 - \gamma$ is connected. We say that a surface S is *non-degenerate* if β_1) for any continuum $\gamma \in G$ the open set $Q_0 - \gamma Q_0$ is connected; β_2) $\gamma \supset Q^*$ for a $\gamma \in G$ implies $\gamma = Q$, i. e., $x(w)$ is constant on Q. The properties α, β_1, β_2 are invariant for Fréchet equivalence. The following statement holds:

LEMMA 3. [1] *Any base surface S whose boundary curve $C = \partial S$ is a Jordan curve is non-degenerate.*

4. We have to deal in the following with subdivisions of Q in rectangles or in triangles. We say that a subdivision of Q in rectangles r is *regular* if the rectangles r are the following ones: $[\alpha_i \leq u \leq \alpha_{i+1}, \beta_j \leq v \leq \beta_{j+1}]$, $i = 0, 1, \cdots, m$, $j = 0, 1, \cdots, n$, where $0 = \alpha_0 < \alpha_1 < \cdots < \alpha_m < \alpha_{m+1} = 1$, $0 = \beta_0 < \beta_1 < \cdots < \beta_n < \beta_{n+1} = 1$, are any given numbers. We say that a subdivision of Q in triangles t is *regular* if each point $w \in Q$ which is a vertex of a triangle t is a vertex also of all adjacent triangles.

We say that a representation $S: x = x(w)$, $w \in Q$, of a surface S is

quasi-linear in Q, or that the vector $x(w)$ is quasi-linear in Q, if $x(w)$ is continuous on Q and there exists a regular subdivision of Q in triangles t such that $x(w)$ is linear on each t. We say that a surface S is a polyhedral surface if S possesses a quasi-linear representation. Each triangle $t \subset Q$ of the corresponding subdivision of Q has for image a triangle $T \subset E_3$ and we call the sum of the elementary areas of the triangles T the elementary area $a(S)$. Given any surface S, we term the Lebesgue area $L(S)$ of S the number $L(S) = \underline{\lim}\, a(\Sigma)$ for all possible polyhedral surfaces Σ when $\| \Sigma, S \| \to 0$.

5. We say that a representation $S: x = x(w), w \in Q$, of a surface S is a D-representation, or that $x(w)$ is a D-vector, if $x(w) = [x^1(u, v), x^2(u, v), x^3(u, v)]$, $w = (u, v) \in Q$, and i) $x^s(u, v)$, $s = 1, 2, 3$, are ACT functions in Q (i. e. absolutely continuous in the sense of Tonelli); ii) the derivatives $x_u{}^s = \partial x^s/\partial u$, $x_v{}^s = \partial x^s/\partial v$, $s = 1, 2, 3$ (which by i) exist a. e. in Q) are L^2-integrable in Q. Every quasi-linear representation is a D-representation.

For any D-vector $x(w)$ let us consider the vectors $x_u = [x_u{}^s, s = 1, 2, 3]$, $x_v = [x_v{}^s, s = 1, 2, 3]$ and the following expressions:

$$E = |\, x_u \,|^2 = \sum_{s=1}^{3} (x_u{}^s)^2, \qquad G = |\, x_v \,|^2 = \sum_{s=1}^{3} (x_v{}^s)^2, \qquad F = x_u \cdot x_v = \sum_{s=1}^{3} x_u{}^s x_v{}^s$$

$$J = J(w) = (J_1, J_2, J_3), \qquad J_s = x_u{}^{s+1} x_v{}^{s+2} - x_v{}^{s+1} x_u{}^{s+2}, \qquad s = 1, 2, 3,$$

where we mean $x^4 = x^1$, $x^5 = x^2$. We call $J = J(w)$ the Jacobian (vector) of the vector $x(w)$. Let us term the *Dirichlet integral* $D[x]$ (or D-integral) and *area integral* $I[x]$ the integrals

$$D[x] = (\tfrac{1}{2}) \int\int_Q (E + G)\, dw,$$

$$I[x] = \int\int_Q |\, J \,|\, dw = \int\int_Q [EG - F^2]^{1/2} dw,$$

where $dw = dudv$. The following statement is well known:

LEMMA 4. [19 and 1, p, 72] *If a surface S possesses a D-representation $S: x = x(w), w \in Q$, then $L(S) = I[x] \leq D[x] < +\infty$.*

We say that a D-representation $S: x = x(w)$, $w \in Q$, of a surface S is *generalized conformal* if iii) $E = G$, $F = 0$ a. e. in Q. For every generalized conformal D-representation we have $(E + G)/2 = [EG - F^2]^{1/2}$ a. e. in Q and $I[x] = D[x]$.

LEMMA 5. [21 and 1, p. 83] *Every non-degenerate surface S with $L(S) < +\infty$ possesses a generalized conformal D-representation $S: x = x(w)$, $w \in Q$. For any representation of this kind we have*

$$L(S) = I[x] = D[x] < +\infty.$$

6. LEMMA 6. *Given a non-degenerate surface S, $L(S) < +\infty$, and any $\epsilon > 0$, there exists a polyhedral surface S_0 having a quasi-linear representation $S_0: x = x_0(w)$, $w \in Q$, such that*

$$\| \partial S, \partial S_0 \| \leq \| S, S_0 \| < \epsilon, \qquad L(S_0) = I[x_0] \leq D[x_0] < L(S) + \epsilon.$$

Proof. By Lemma 5 there exists a generalized conformal D-representation $S: x = x(w)$, $w \in Q$, of S and $L(S) = I[x] = D[x] < +\infty$. Let $x(w) = [x^s(u, v), s = 1, 2, 3]$, $w = (u, v) \in Q$, and let $x_n{}^s(u, v)$, $s = 1, 2, 3$, be the Stieltjes polynomials of order n of the continuous functions $x^s(u, v)$, $s = 1, 2, 3$ (the latter supposed defined in the square $(-1 \leq u \leq 1$, $-1 \leq v \leq 1)$ by symmetry and in the whole plane (u, v) by periodicity with period 2 with respect to u and v). Let S_n be the elementary surfaces $S: x = x_n(w)$, $w \in Q$, $n = 1, 2, \cdots$. We have (Lemma 4) $L(S_n) = I[x_n] \leq D[x_n] < +\infty$. Because the functions $x_n{}^s(u, v)$ are continuous in E_2, are ACT in each finite region and their derivatives $(x_n{}^s)_u$, $(x_n{}^s)_v$ are L^2-integrable in Q, then, by well known theorems [9, 27], we have $x_n(w) \rightrightarrows x(w)$ on Q and the integrals of $[EG - F^2]^{1/2}$ and of $E + G$ in Q, calculated with the polynomials $x_n{}^s(u, v)$, approach the corresponding integrals calculated with the functions $x^s(u, v)$, $s = 1, 2, 3$; that is, $I[x_n] \to I[x]$, $D[x_n] \to D[x]$ as $n \to \infty$. Hence there is an integer n such that $|x_n(w) - x(w)| < \epsilon/2$ for any $w \in Q$, and

$$L(S_n) = I[x_n] \leq D[x_n] < D[x] + \epsilon/2 = L(S) + \epsilon/2.$$

Now let us consider the surface S_n in the representation $S: x = x_n(w)$, $w \in Q$, $x_n(w) = [x_n{}^s(u, v), s = 1, 2, 3]$, where the $x_n{}^s$ are polynomials. A well known elementary procedure permits us to define a polyhedral surface S_0 (inscribed in S_n) and a quasi-linear representation $S_0: x = x_0(w)$, $w \in Q$, of S_0 such that $|x_n(w) - x_0(w)| < \epsilon/2$ for any $w \in Q$ and such that $|I[x_0] - I[x_n]| < \epsilon/2$, $|D[x_0] - D[x_n]| < \epsilon/2$. Therefore $|x(w) - x_0(w)| < \epsilon/2 + \epsilon/2 = \epsilon$ for any $w \in Q$ and (Lemma 5) $L(S_0) = I[x_0] \leq D[x_0] < D[x_n] + \epsilon/2 < L(S) + \epsilon$. Finally by Lemma 2 we have

$$\| \partial S, \partial S_0 \| \leq \| S, S_0 \| \leq \max |x(w) - x_0(w)| < \epsilon.$$

Lemma 6 is thereby proved.

7. Let $A \subset E_3$ be a closed subset of E_3; let $F(x, p)$, $x = (x^1, x^2, x^3)$, $p = (p^1, p^2, p^3)$, be any single-valued real function of the real vectors x, p such that

a) $F(x, p)$ is continuous as a function of (x, p) for any $x \in A$ and $|p| \neq 0$;

b) $F(x, tp) = tF(x, p)$ for any $x \in A$, $|p| \neq 0$, $t > 0$.

Let S be any surface such that $[S] \subset A$, $L(S) < +\infty$. Let $S: x = x(w)$, $w \in Q$, be any representation of S. Then a concept of Weierstrasse integral on the surface S, $\mathfrak{J}(S) = \int\int_S F(x, p)dw$ $(dw = dudv)$, has been previously introduced [2], and $\mathfrak{J}(S)$ is invariant for Fréchet equivalence. If $F(x, p) = |p|$, then $\mathfrak{J}(S) = L(S)$. The following statements hold:

LEMMA 7. *If S possesses the D-representation $S: x = x(w)$, $w \in Q$, then $\mathfrak{J}(S)$ is given by the classical Lebesgue integral*

$$\mathfrak{J}(S) = \int\int_S F(x, p)dw = \int\int_Q F[x(w), J(w)]dw,$$

where $J(w) = (J_1, J_2, J_3)$ is the Jacobian (vector) of the vector $x(w)$ (no. 5).

LEMMA 8. *If S, S_n, $n = 1, 2, \cdots$, are surfaces of finite Lebesgue area and $[S], [S_n] \subset A$, $\| S_n, S \| \to 0$, $L(S_n) \to L(S)$, then $\mathfrak{J}(S_n) \to \mathfrak{J}(S)$.*

Lemma 7 is a particular case of a more general statement [2, p. 107]. Lemma 8 has been proved in [2, p. 101].

8. We say that $\mathfrak{J}(S)$ is a *positive definite* (semidefinite) integral if $F(x, p) > 0$ $(F(x, p) \geq 0)$ for any $x \in A$, $|p| \neq 0$. Let us now suppose that

c) the derivatives $F_s(x, p) = \partial F(x^1, x^2, x^3, p^1, p^2, p^3)/\partial p^s$, $s = 1, 2, 3$, exist and are continuous for any $x \in A$, $|p| \neq 0$.

We say that $\mathfrak{J}(S)$ is a *positive regular* (semiregular) integral if

$$\mathfrak{E}(x, p, \bar{p}) = F(x, \bar{p}) - \Sigma \bar{p}^s F_s(x, p) > 0 \ (\geq 0)$$

for any $x \in A$, $|p|$, $|\bar{p}| \neq 0$, $p \neq \bar{p}$. We say that $\mathfrak{J}(S)$ is a lower semicontinuous integral on a surface S_0 with respect to a family $\{S\}$ of surfaces S if, given $\epsilon > 0$, there exists a $\delta > 0$ such that for any surface $S \in \{S\}$, $\| S, S_0 \| < \delta$, we have $\mathfrak{J}(S) > \mathfrak{J}(S_0) - \epsilon$. The following statement has been previously proved:

LEMMA 9. *If $F(x, p) > 0$ for any $x \in A$, $|p| \neq 0$, if S_0 is a given surface and $[S_0] \subset A$, $L(S_0) < +\infty$, and if $\mathfrak{E}(x, p, \bar{p}) \geq 0$ for any $x \in [S_0]$,*

$p \neq \bar{p}$, $|p|$, $|\bar{p}| \neq 0$, then the integral $\mathfrak{J}(S)$ is lower semicontinuous on the surface S_0 with respect to all surfaces S such that $[S] \subset A$, $L(S) < +\infty$.

This lemma implies that any positive definite semiregular integral $\mathfrak{J}(S)$ is lower semicontinuous on each surface S_0 such that $[S_0] \subset A$, $L(S_0) < +\infty$. Lemma 9, as well as this last statement, are particular cases of a previous theorem [3, p. 67].

Note. For the validity of Lemma 9 it is sufficient to suppose that condition c) holds for all $x \in [S_0]$ (not necessarily for all $x \in A$), as can be easily verified by inspecting the mentioned proof given in [3], especially pp. 61-68.

9. LEMMA 10. *To any surface S we can associate a base surface S_0 having the same boundary curve $\partial S_0 = \partial S$ and such that $[S_0] \subset [S]$, $L(S_0) \leq L(S)$. In addition, if $\mathfrak{J}(S)$ is any positive semidefinite integral and $[S] \subset A$, $L(S) < +\infty$, then we have also $\mathfrak{J}(S_0) \leq \mathfrak{J}(S)$.*

LEMMA 11. *Given a Jordan curve C of E_3, given $\epsilon > 0$, there exists a $\delta = \delta(C, \epsilon) > 0$ such that to each base surface S with $\| \partial S, C \| < \delta$ we can associate a non-degenerate surface S_0, such that $[S_0] \subset [S]$, $L(S_0) \leq L(S)$, $\| \partial S_0, C \| < \epsilon$. In addition, if $\mathfrak{J}(S)$ is any positive semidefinite integral and $[S] \subset A$, $L(S) < +\infty$, then we have also $\mathfrak{J}(S_0) \leq \mathfrak{J}(S)$.*

Lemma 10 is a result which is well known and often applied [16, 20, 7]; Lemma 11 has been recently proved [7]. Both results depend on the so-called retraction process for surfaces [7].

10. *We say that a given sequence $S_n : x = x_n(w)$, $w \in Q$, $n = 1, 2, \cdots$,* of representations of surfaces, that is, a given sequence of vectors, satisfies a *three points condition* on Q^* if Q^* contains for each n three points a_n, b_n, c_n, whose mutual distances exceed some fixed $\delta > 0$ and such that the three points $x_n(a_n)$, $x_n(b_n)$, $x_n(c_n)$ tend to three distinct limits as $n \to \infty$.

LEMMA 12. *If S_n, $n = 1, 2, \cdots$, are non-degenrate surfaces such that $L(S_n) < +\infty$ and $\| \partial S_n, C \| \to 0$, where C is a continuous closed curve not reduced to a single point, then there are generalized conformal D-representations $S_n : x = x_n(w)$, $w \in Q$, of the surfaces S_n, $n = 1, 2, \cdots$, satisfying a three points condition on Q^*.*

This Lemma is only a variant of Lemma 5 [1, 21, 35].

11. LEMMA 13. *Given any two positive numbers N, ϵ, there exists a positive number $\eta = \eta(N, \epsilon) > 0$ with the following property: To any D-vector $x(w)$, $w \in Q$, such that $D[x] < N$ we can associate a number $\delta = \delta[N, \epsilon, x(w)] > 0$, $\eta < \delta < \epsilon$, and a regular subdivision* (no. 4) *of Q in rectangles r whose dimensions are between δ and 2δ and such that the image of each side (not on Q^*) of the rectangles r on the surface $S: x = x(w)$, $w \in Q$, is a rectifiable curve whose length is $< \epsilon$.*

This Lemma has been proved recently [35, especially the first pages, pp. 317-319].

12. LEMMA 14. *Given a sequence of D-representations of surfaces $S_n: x = x_n(w)$, $w \in Q$, $n = 1, 2, \cdots$, such that $D[x_n] < N$, N constant, $n = 1, 2, \cdots$, and $\| \partial S_n, C \| \to 0$, where C is a closed Jordan curve, then the vectors $x_n(w)$, $n = 1, 2, \cdots$, are equicontinuous on Q^*.*

This Lemma has been proved in different ways [1, 21, 35].

§ 2. A Smoothing Process for Polyhedral Surfaces.

13. LEMMA 15. *Suppose $f(x)$ is L-integrable and non-negative on $0 \leq x \leq a$, and that $0 < \int_x^a f(\alpha)\,d\alpha \leq kf^2(x)$ almost everywhere on $0 \leq x < a$ and in particular for $x = 0$. Then $a \leq 2kf(0)$.*

Proof. The transformation $g = kf$ shows we may assume $k = 1$. Set $h(x) = \int_x^a f(\alpha)\,d\alpha$. Then $h(x) > 0$ for all $0 \leq x < a$ and $h'(x) = -f(x)$ a.e.; hence $0 < h(x) \leq [h'(x)]^2$ a. e., $h'(x) \leq -\sqrt{h(x)} < 0$ a. e., and and $h(x) \geq \int_x^a \sqrt{h(\alpha)}\,d\alpha$ on $0 \leq x \leq a$. Let $h_1(x) = \int_x^a \sqrt{h(\alpha)}\,d\alpha$, $h_2(x) = \int_x^a \sqrt{h_1(\alpha)}\,d\alpha$, $0 \leq x \leq a$. Then $h \geq h_1 \geq h_2$ on $0 \leq x \leq a$, and $h, h_1, h_2, -h_1', -h_2'$ are continuous on $0 \leq x \leq a$, and positive on $0 \leq x < a$; h_2'' also is continuous on $0 \leq x < a$, and $h_2''(x) = (1/2)\sqrt{h(x)}/h_1(x) \geq 1/2$ on $0 \leq x < a$. Hence, by the theorem of the mean, $-h_2'(x) \geq (a - x)/2$ and so $h_2(x) \geq (a - x)^2/4$, $f^2(0) \geq h(0) \geq h_2(0) \geq a^2/4$.

14. LEMMA 16. *If C is a simple closed polygonal in E_3, l the length of C, $\pi \subset E_2$ a simple polygon, $C: x = x(w)$, $w \in \pi^*$, a quasi-linear representation of C on π^*, then there is a polyhedral surface S whose boundary*

curve is $\partial S = C$ and a quasi-linear representation $S: x = X(w)$, $w \in \pi$, of S on π, such that $X(w) = x(w)$ for any $w \in \pi^*$, $a(S) \leq l^2/4$, and $[S]$ is contained in the minimum convex body containing C.

Proof. Let w_i, $i = 1, 2, \cdots, n$, $w_{n+1} = w_1$ be points of π^* such that $x(w)$ is linear on each arc $\tau_i = w_i w_{i+1}$, $i = 1, 2, \cdots, n$, of π^*. Let $x_i = x(w_i)$, $x_{n+1} = x_1$, $i = 1, 2, \cdots, n$, be the images of w_i on C. We can say that the x_i are vertices of C. Let S be the polyhedral surface sum of the triangles $T_i = x_1 x_i x_{i+1}$, $i = 2, 3, \cdots, n-1$. If d is the diameter of C, we have $d \leq l/2$ and

$$\begin{aligned} \text{area } T_i &\leq (1/2)\,|\,x_1 - x_i\,|\,|\,x_i - x_{i+1}\,| \leq (1/2)d\,|\,x_i - x_{i+1}\,| \\ &\leq (l/4)\,|\,x_i - x_{i+1}\,|, \qquad\qquad i = 2, 3, \cdots, n-1. \end{aligned}$$

Finally area $S \leq (l/4)\Sigma\,|\,x_i - x_{i-1}\,| \leq l^2/4$. Let $w' = h(w)$, $w' = (u', v')$, be a quasi-linear homeomorphism between π and the triangle

$$T \equiv [u' \geq 0, v' \geq 0, u' + v' \leq 1]$$

such that, if $w_i' = h(w_i)$ are the images of w_i on T^*, $i = 1, 2, \cdots, n$, we have $w_1' = (0, 0)$, $w_2' = (1, 0)$, $w_n' = (0, 1)$. All other points w_i', $i = 3, 4, \cdots, n-1$, are on the segment $(u' \geq 0, v' \geq 0, u' + v' = 1)$. We can also suppose that $h(w)$ is linear on each τ_i, $i = 1, 2, \cdots, n$. Let $X'(w')$ be the quasi-linear continuous vector, $w' \in T$, which is linear on each triangle $w_1' w_i' w_{i+1}'$ and there represents the triangle T_i, $i = 2, 3, \cdots, n-1$, and is such that $X'(w_i') = x_i$, $i = 1, 2, \cdots, n$. Let finally $X(w) = X'[h(w)]$, $w \in \pi$. The vector $X(w)$, $w \in \pi$, has all the required properties.

15. Let S be a polyhedral surface, $S: x = x(w)$, $w \in Q$, a quasi-linear representation of S on Q, $C = \partial S$ the boundary curve of S. Let x_0 be any point of E_3, $\rho > 0$ any real number, $B^* = B^*(\rho) \equiv [\,|\,x - x_0\,| = \rho]$ the sphere of centre x_0 and radius ρ, $B = B(\rho) \equiv [\,|\,x - x_0\,| < \rho]$ the interior of B^*. Let us suppose $C \subset B(r)$ for a given $r > 0$, and, for any $\rho \geq r$, let us consider the maximal connected subset $A = A(\rho)$ of Q containing Q^* and open with respect to Q, such that $x(w) \in B(\rho)$ for any $w \in A(\rho)$. Let R be the maximum distance of the points of $[S]$ from x_0. We have $R \geq r$.

LEMMA 17. *Under the above conditions $A(\rho)$ is an open connected set whose boundary is $Q^* + \Sigma\pi_i^*$, where π_i, $i = 1, 2, \cdots, n$, are simple disjoint Jordan regions, finite in number, and each π_i^* is a sum of elementary arcs. If $\mathfrak{Q}$ is the image of $\Sigma\pi_i^*$ on S, then $\mathfrak{Q} \subset B^*(\rho)$ is a sum of a finite number*

of elementary curves and, if $l(\rho)$ is the total length of $\mathfrak{L}$, we have $l(\rho-0)=l(\rho)\leq l(\rho+0)$ for any $r\leq\rho\leq R$.

Proof. The vector $x(w)$ is quasi-linear on Q, hence is linear on the triangles of a regular subdivision of Q in triangles. Let v, d, t denote the vertices, sides, triangles of the subdivision. Let V, D, T denote the images of the v, d, t. Let us consider the distances of all elements V, D, T from x_0 (in the sense of set theory, no. 1). We have a finite collection $\{\rho\}$ of numbers ρ. All elements V are points of E_3, all elements D are segments or points, all elements T are triangles, or segments, or points. For each triangle T the intersection TB^* is one (or two, or three) arcs λ of circumference (or a single point). Hence $\mathfrak{L}$, as a part of the intersection SB^*, is a sum of a finite number of closed curves p_i, $i=1,2,\cdots,n$, each a sum of arcs of circumference λ. This holds for any p, $r\leq\rho\leq R$. Let us denote by $\mathfrak{A}=\mathfrak{A}(\rho)$ the image of $A=A(\rho)$. Each arc λ is the image of an arc of a conic in a triangle t in Q and the curves π_i^* are sums of these arcs, plus, possibly, segments, finite in number, whose images are single points which are intersections of segments T with B^*. If ρ is not in the finite collection $\{\rho\}$, no point V is on B^*, hence for each arc λ the vertices of the triangle T to which λ belongs are not on B^*. Hence λ separates two annular strips in T, one in $\mathfrak{A}$ and in B, one not in $\mathfrak{A}$ and outside of $B+B^*$. The curves ρ_i and their total length $l(\rho)$ vary continuously as ρ varies in a sufficiently small neighborhood of the given ρ, i. e. $l(\rho-0)=l(\rho)=l(\rho+0)$.

If ρ is a number in the collection $\{\rho\}$, then suppose first that there is a vertex on B^*. If V is such a vertex, for each triangle T having a vertex in V let us consider the part of T contained in a circle of center V and sufficiently small radius. This part may be completely contained in B, or completely outside of $B+B^*$, or partially in B and partially outside of $B+B^*$; and in this last case such a part contains an arc of circumference λ ending in V which separates points in B from points outside of $B+B^*$. Therefore the neighborhood u of V on S can be divided in sectors τ alternatively in B and outside of $B+B^*$. Not all sectors τ which are in B belong to $\mathfrak{A}$. If such a sector τ is in B and belongs to $\mathfrak{A}$, then the two corresponding arcs λ ending in V are adjacent arcs on a curve p_i. If a sector τ is in B but does not belong to $\mathfrak{A}$, then τ is a part of a region Σ in S which is completely inside of B and which belongs to $\mathfrak{A}(\rho')$ for all $\rho'>\rho$. As ρ varies continuously in a left neighborhood of the given value ρ, all arcs λ vary continuously, remaining inside of $\mathfrak{A}$; hence $l(\rho-0)=l(\rho)$. As ρ varies continuously in a right neighborhood of the given value ρ, again all arcs λ

vary continuously, but we have new arcs λ, which approach the arcs λ of the boundaries of the regions Σ. This assures that $l(\rho+0)$ exists and $l(\rho+0) \geq l(\rho)$.

Analogous reasoning holds for each of the values $\rho \in \{\rho\}$ which equal the distances of segments D, or triangles T, from x_0.

We have proved that $l(\rho-0) = l(\rho) = l(\rho+0)$ for any $r \leq \rho \leq R$, ρ not $\in \{\rho\}$; $l(\rho-0) = l(\rho) \leq l(\rho+0)$ for any $r \leq \rho \leq R$, $\rho \in \{\rho\}$. Lemma 17 is thereby proved.

16. With the same notation as in no. 15 let $\sigma(\rho)$ be the collection of the elementary surfaces defined by the vector $x(w)$ on the Jordan regions π_i, $i = 1, 2, \cdots, n$. Let $\alpha(\rho)$ be the total area of $\sigma(\rho)$. Evidently $\alpha(R) = 0$, $\alpha(\rho) \leq a(S)$, $r \leq \rho \leq R$, and $\alpha(\rho)$ is a monotone non-increasing function of ρ, $r \leq \rho \leq R$. In addition, the previous discussion (proof of Lemma 17) assures us that $\alpha(\rho)$ is continuous at each point ρ, $r \leq \rho \leq R$, ρ not $\in \{\rho\}$.

LEMMA 18. *Under the above conditions we have, for any* ρ,

$$\alpha(\rho) \geq \int_{\rho}^{R} l(\rho)\, d\rho, \qquad r \leq \rho \leq R.$$

Note. This Lemma has a quite evident and elementary proof which we are going to give in the following lines (see also 32). For a higher formulation of this Lemma, not necessary for our purpose, see [12, 13].

Proof of Lemma 18. First of all let us observe that $a(\mathfrak{A}) + a(\sigma) = a(S)$. Now let us consider any two values of ρ, $\rho_1 < \rho_2$, such that the interval (ρ_1, ρ_2) is free of points of $\{\rho\}$; then the difference $\alpha(\rho_1) - \alpha(\rho_2)$ is the sum of a finite number of the elementary areas described by the arcs of circumference $\lambda = \lambda(\rho)$ on the corresponding triangles T of S. If we consider one of these T and we call h the distance of x_0 from the plane of T, x_1 the projection of x_0 on such a plane, then $\lambda = \lambda(\rho)$ is an arc of circumference of center x_1 and radius $[\rho^2 - h^2]^{1/2}$. If we indicate with $\lambda(\rho)$ also the length of $\lambda(\rho)$ [the endpoints of $\lambda(\rho)$ are on the boundary of T], then, as ρ varies between ρ_1 and ρ_2, $\lambda(\rho)$ describes the elementary area

$$\tau = \int_{\rho_1}^{\rho_2} \lambda(\rho)\, d[\rho^2 - h^2]^{1/2} = \int_{\rho_1}^{\rho_2} \lambda(\rho)\, \{\rho/[\rho^2 - h^2]^{1/2}\}\, d\rho \geq \int_{\rho_1}^{\rho_2} \lambda(\rho)\, d\rho.$$

By adding these relations for all triangles T containing an arc λ we have

$$\alpha(\rho_1) - \alpha(\rho_2) \geq \int_{\rho_1}^{\rho_2} l(\rho)\, d\rho. \tag{1}$$

Let us suppose that ρ, $r \leq \rho \leq R$, be any given number and that $\rho \leq \rho_1 < \rho_2 < \cdots < \rho_m \leq R$ are all the distinct members of $\{\rho\}$ between ρ and R. By (1), which holds on each interval interior to (ρ_i, ρ_{i+1}), we have

$$\alpha(\rho_i + 0) - \alpha(\rho_{i+1} - 0) \geq \int_{\rho_i}^{\rho_{i+1}} l(\rho)\, d\rho, \qquad i = 1, 2, \cdots, m-1,$$

and, $\alpha(\rho)$ being monotone non-increasing, also

$$\alpha(\rho_i) - \alpha(\rho_{i+1}) \geq \int_{\rho_i}^{\rho_{i+1}} l(\rho)\, d\rho, \qquad i = 1, 2, \cdots, m-1.$$

Analogously

$$\alpha(\rho) - \alpha(\rho_1) \geq \int_{\rho}^{\rho_1} l(\rho)\, d\rho, \qquad \alpha(\rho_m) - \alpha(R) \geq \int_{\rho_m}^{R} l(\rho)\, d\rho,$$

if $\rho < \rho_1$, or $\rho_m < R$. In any case $\alpha(R) = 0$, hence, by adding,

$$\alpha(\rho) \geq \int_{\rho}^{R} l(\rho)\, d\rho.$$

17. We will prove, in nos. 19-24, the following

LEMMA 19. *Let S be a polyhedral surface, $S: x = x(w)$, $w \in Q$, a quasi-linear representation of S on Q, $C = \partial S$ the boundary curve of S, L the length of C, D any number $\geq$ the diameter of C, x_0 any point of C, K any positive constant. Then there exists a polyhedral surface S_0 and a quasi-linear representation $S_0: x = x_0(w)$, $w \in Q$, of S_0 on Q, with the following properties:*

i) *S_0 is contained in a sphere of center x_0 and radius*

$$\rho \leq 2D + 3[Ka(S)]^{1/2}.$$

ii) *there is an open set $\pi \subset Q$, which is the sum $\Sigma\pi_i$ of a finite number of disjoint simple polygons π_i, $i = 1, 2, \cdots, n$, such that $x_0(w) = x(w)$ on $Q - \pi$; hence $x_0(w) = x(w)$ on each π_i^*, $i = 1, 2, \cdots, n$, and on Q^*. Denote by σ, σ_0 $[\sigma_i, \sigma_{0i}]$ the polyhedral surfaces represented by $x(w)$, $x_0(w)$ on π $[\pi_i, i = 1, 2, \cdots, n]$ and by $a(\sigma)$, $a(\sigma_0)$ their total area, $\sigma = \Sigma\sigma_i$, $\sigma_0 = \Sigma\sigma_{0i}$, $a(\sigma) = \Sigma a(\sigma_i)$, $a(\sigma_0) = \Sigma a(\sigma_{0i})$. Each surface σ_{0i} is contained in the minimum convex body containing the curve $\partial\sigma_i = \partial\sigma_{0i}$. Denote by p_i the curves $\partial\sigma_i = \partial\sigma_{0i}$, and also their lengths, $i = 1, 2, \cdots, n$; denote by $\mathfrak{L}$ the family of curves p_i, $i = 1, 2, \cdots, n$, and also their total length, $\mathfrak{L} = \Sigma p_i$; then*

iii) $\mathfrak{L}^2 = (\Sigma p_i)^2 \leq K^{-1} a(\sigma)$;

iv) $a(\sigma_0) \leq (K^{-1}/4) a(\sigma)$;

v) *each continuum $c \subset Q$ such that $cQ^* \neq 0$, diam $x(c) < D$, is completely contained in $Q - \pi$.*

18. *Note.* 1) If we suppose D exactly the diameter of C, then $D \leq L/2$ and the inequality i) becomes:

$$\rho \leq L + 3[Ka(S)]^{1/2}.$$

2) In the statements of Lemma 19 it is not excluded that π is empty, hence $S = S_0$, $\mathfrak{L} = 0$, $a(\sigma_0) = 0$.

19. The curve C is completely interior to any sphere $B^*(r)$ of center x_0 and radius $r > D$. We can take for r the value $r = 2D$. Let us call R the maximum distance of the points x of $[S]$ from x_0. Then S is contained in the sphere $B^*(R)$ of center x_0 and radius R. If $R \leq 2D + 3[Ka(S)]^{1/2}$ the Lemma is proved. Let us suppose therefore $R > 2D + 3[Ka(S)]^{1/2}$, hence $R > r$. Let us consider for any ρ, $r \leq \rho \leq R$, the set $A = A(\rho)$ studied in nos. 15, 16, the functions $l(\rho) \geq 0$, $\alpha(\rho) \geq 0$ $r \leq \rho \leq R$, and the finite family $\sigma = \sigma(\rho)$ of elementary surfaces contained in S and represented by the vector $x(w)$ on the finite family $\pi = \Sigma\pi_i$ of simple Jordan regions, $\pi = Q - (A + A^*)$ (no. 15). We know (Lemma 17) that $l(\rho - 0) = l(\rho) \leq l(\rho + 0)$ for any $r \leq \rho \leq R$ and that $l(\rho)$ is the total length of the boundary of σ and $\alpha(\rho)$ the total area of σ. In addition (Lemma 18), for any $r \leq \rho \leq R$, we have

$$\int_\rho^R l(\alpha)\, d\alpha \leq \alpha(\rho).$$

Let I be the set of all ρ such that $l(\rho) = 0$. The set I is not empty because $l(R) = 0$. If ρ_0 is a limit point at the left for I then $l(\rho_0 - 0) = l(\rho_0) = 0$; if ρ_0 is a limit point at the right for I, then $0 \leq l(\rho_0 - 0) = l(\rho_0) \leq l(\rho_0 + 0) = 0$, hence $l(\rho_0) = 0$. This proves that I is closed. Let us indicate by R_0, $r \leq \rho \leq R$, the smallest $\rho \in I$, $r \leq \rho \leq R$. Let us suppose $R_0 \leq 2D + 3[Ka(S)]^{1/2}$. We have $l(R_0) = 0$, i. e. the vector $x(w)$ represents on π^* curves reduced to a single point, hence $x(w)$ is constant on each simple closed curve π_i^* bounding a component π_i of π. Let us define the continuous vector $x_0(w)$ on Q by saying that $x_0(w)$ coincides with $x(w)$ in $Q - \pi$, and $x_0(w)$ is constant on each component π_i of π. We have $a(\sigma_0) = 0$, $\mathfrak{L} = 0$, and the surface S_0 is contained in the sphere $B^*(R_0)$. Because the vector $x(w)$ is quasi-linear in Q, the simple closed curves π_i^* bounding the components of π are sums of a finite number of segments, that is the components of π are simple polygons in Q; and this assures also that the vector $x_0(w)$ is quasi-linear on Q. Lemma 19 is therefore proved if $R_0 \leq 2D + 3[Ka(S)]^{1/2}$.

20. Let us suppose now $R_0 > 2D + 3[Ka(S)]^{1/2}$, hence $r < R_0 \leq R$. We have $l(\rho) > 0$ for any $r \leq \rho < R_0$. Let us observe that the interval $(r, r + [Ka(S)]^{1/2})$ is completely contained in the interval (r, R_0). Let us suppose, if possible, that $l(\rho) > [K^{-1}a(S)]^{1/2}$ for any $r \leq \rho \leq r + [Ka(S)]^{1/2}$. Then by Lemma 18 we would have

$$a(S) > \alpha(r) \geq \int_r^R l(\rho)\,d\rho > \int_r^{r+[Ka(S)]^{\frac{1}{2}}} [K^{-1}a(S)]^{1/2}\,d\rho = a(S),$$

which is contradictory. Hence there must be at least one number r', $r \leq r' \leq r + [Ka(S)]^{1/2}$, such that $l(r') \leq [K^{-1}a(S)]^{1/2}$. (For the reasoning in this number see [32].)

21. Let us consider the functions of ρ, $\phi(\rho) = \int_\rho^{R_0} l(\alpha)\,d\alpha$, $Kl^2(\rho)$, $r' \leqq \rho \leqq R_0$, where $\phi(\rho)$ is continuous and $l(\rho - 0) = l(\rho) \leq l(\rho + 0)$ for any ρ. Let J be the set of all numbers ρ, $r' \leq \rho \leq R_0$, such that $\phi(\rho) > Kl^2(\rho)$ and let r'' be the g. l. b. of the numbers $\rho \in J$. Let $r'' = R_0$ if J is empty.

Let us observe that for $\rho = R_0$ we have $\phi(R_0) = 0$, $l(R_0) = 0$; hence $\rho = R_0$ does not belong to J. In consequence we have $r'' = R_0$ only if J is empty.

A) If $r'' = r'$ then either $r' \in J$ and $\phi(r') > Kl^2(r')$, or there are $\rho \in J$ as close as we want to r' for which $\phi(\rho) > Kl^2(\rho)$, hence $\phi(r') \geq Kl^2(r' + 0) \geq Kl^2(r')$. In both cases

$$\int_{r''}^{R_0} l(\alpha)\,d\alpha \geq Kl^2(r''), \qquad r'' = r'.$$

B) If $r' < r'' < R_0$, then we have $\phi(\rho) \leq Kl^2(\rho)$ for any $r' \leq \rho < r''$; hence $\phi(r'') \leq Kl^2(r'' - 0) = Kl^2(r'')$. On the other hand there are numbers $\rho \in J$, $r' < \rho < R_0$, as close as we want to r'' for which $\phi(\rho) > Kl^2(\rho)$; hence $\phi(r'') \geq Kl^2(r'' + 0)$. It follows that $l(r'' - 0) \geq l(r'' + 0)$, while we know that $l(r'' - 0) = l(r'') \leq l(r'' + 0)$. Therefore $l(r'' - 0) = l(r'') = l(r'' + 0)$ and $\phi(r'') = Kl^2(r'')$. We conclude that

$$\int_\rho^{r''} l(\alpha)\,d\alpha \leq \int_\rho^{R_0} l(\alpha)\,d\alpha \leq Kl^2(\rho)$$

for any $r' \leq \rho \leq r''$, where $l(\rho) > 0$ for any $r' \leq \rho \leq r''$ and

$$\int_{r''}^{R_0} l(\alpha)\,d\alpha = Kl^2(r'').$$

By Lemma 15 it follows that

$$r'' - r' \leq 2Kl(r') \leq 2K[K^{-1}a(S)]^{1/2} = 2[Ka(S)]^{1/2}.$$

C) If $r'' = R_0$ then J is empty, $\phi(R_0) = 0$, $l(R_0) = 0$, $\int_{\rho}^{R_0} l(\alpha)\,d\alpha \leq Kl^2(\rho)$ for any $r' \leq \rho \leq R_0$, $l(\rho) > 0$ for any $r' \leq \rho < R_0$, and finally $0 = \int_{R_0}^{R_0} l(\alpha)\,d\alpha = Kl^2(R_0)$. Again, by Lemma 15, we have

$$r'' - r' = R_0 - r' \leq 2Kl(r') \leq 2[Ka(S)]^{1/2}.$$

Let us observe that in all three cases we have

$$r'' \leq r' + 2[Ka(S)]^{1/2} \leq r + 3[Ka(S)]^{1/2} = 2D + 3[Ka(S)]^{1/2},$$

$$\int_{r''}^{R_0} l(\alpha)\,d\alpha \geq Kl^2(r'').$$

22. From no. 15 we know that the boundary $\mathfrak{L}$ of the surface $\sigma = \sigma(r'')$ is the sum of a finite number of closed elementary curves $p_1, p_2, \cdots, p_n$ and that each p_j is a sum of a finite number of arcs of circumference $\lambda = xx'$ of the sphere $B^*(r'')$. Each λ is contained in a triangle T of S and has its end point x, x' on the boundary T^* of T. Hence the plane region between the arc $\lambda = xx'$ and the segment $\lambda' = (xx')$ (both with endpoints x, x' on T^*) is completely contained in T, is interior to the sphere $B^*(r'')$ and belongs to the part $\mathfrak{A}(\rho)$ of the surface S. If we substitute each arc λ of the curve p_j with the segments λ' we obtain a polygonal line p_j' inscribed in p_j and completely contained in the sphere $B^*(r'')$. For each triangle T (containing an arc λ, let us consider the corresponding triangle t of Q; then λ and λ' have counterimages in t which are respectively an arc τ of a conic and a segment τ' having the same endpoints on the boundary t^* of t, and the plane region between τ and τ' belongs to t and to $A = A(r'')$, whereas τ belongs to $A^* = A^*(r'')$. Then p_j' is the image under the vector $x(w)$ of the polygonal line $\pi_j'^*$ which we obtain by substituting the arcs τ of A^* with the segments τ', $p_j' : x = x(w)$, $w \in \pi_j'^*$. This holds for any $j = 1, 2, \cdots, n$; hence we have, on Q, n simple polygons π_j', $j = 1, 2, \cdots, n$, without common interior points and, if π' is their sum, $Q - \pi' \subset A = A(r'')$, $\pi' = \Sigma\pi_j'$, $\pi_j' \supset \pi_j'$, $j = 1, \cdots, n$. The polygonal closed curves p_j', $j = 1, 2, \cdots, n$, are all contained in the sphere $B^*(r'')$ and we indicate their sum by $\mathfrak{L}'$. If $\mathfrak{L}'$ is also the total length of $\mathfrak{L}'$, we have $\mathfrak{L}' \leq l(r'')$, where $l(r'')$ is the total length of $\mathfrak{L}$. Let us call p_j' also the length of p_j', $j = 1, 2, \cdots, n$; hence $p_1' + p_2' + \cdots + p_n' = \mathfrak{L}'$. Let us denote by σ_j' the surfaces defined by the vector $x(w)$ on π_j and by σ' their sum. We have $a(\sigma') \geq a(\sigma)$, where $\sigma = \sigma(r'')$ and $a(\sigma) = \alpha(r'')$; hence $\alpha(r'') \leq a(\sigma')$.

23. For each polygonal closed curve p_j' there is (Lemma 16) a polyhedral surface σ_{0j} having p_j' as a boundary curve, and a quasi-linear representation $\sigma_{0j}: x = x_j(w)$, $w \in \bar{\pi}_j' = \pi_j' + \pi_j'^*$, such that $x_j(w) = x(w)$ for any $w \in \pi_j'^*$ and $a(\sigma_{0j}) \leq (1/4)p_j^2$, $j = 1, 2, \cdots, n$. Let us denote by σ_0 the finite family of surfaces σ_{0j}, $j = 1, 2, \cdots, n$, and by $x_0(w)$, $w \in Q$, the continuous vector defined by $x_0(w) = x_j(w)$ for any $w \in \pi_j'$, $j = 1, 2, \cdots, n$, $x_0(w) = x(w)$ for any $w \in Q - \pi'$. The vector $x_0(w)$ is quasi-linear on Q, hence the surface $S_0: x = x_0(w)$, $w \in Q$, is a polyhedral surface. The polygonal closed curves p_j', $j = 1, 2, \cdots, n$, are contained in the sphere $B^*(r'')$ (the vertices are points on the sphere $B^*(r'')$). By Lemma 16 each surface σ_{0j} is contained in the minimum convex body containing p_j', hence each surface σ_{0j}, and consequently also the whole surface S_0, is contained in the sphere $B^*(r'')$. Thus i) is proved.

24. By Lemma 16 and all the preceding considerations, we have

$$\begin{aligned} a(\sigma_0) = \Sigma a(\sigma_{0j}) &\leq (1/4)\Sigma p_j'^2 \leq (1/4)(\Sigma p_j')^2 = (1/4)\mathfrak{L}'^2 \\ &\leq (1/4)l^2(r'') \leq (K^{-1}/4)\int_{r''}^{R_0} l(\alpha)\,d\alpha \leq (K^{-1}/4)\int_{r''}^{R_0} l(\alpha)\,d\alpha \\ &\leq (K^{-1}/4)\alpha(r'') \leq (K^{-1}/4)a(\sigma'); \end{aligned}$$

that is, $a(\sigma_0) \leq (K^{-1}/4)a(\sigma')$, and iv) is proved. We have

$$\begin{aligned} \mathfrak{L}'^2 = (\Sigma p_j')^2 \leq l^2(r'') &\leq K^{-1}\int_{r''}^{R_0} l(\alpha)\,d\alpha \leq K^{-1}\int_{r''}^{R} l(\alpha)\,d\alpha \\ &\leq K^{-1}\alpha(r'') \leq K^{-1}a(\sigma'); \end{aligned}$$

thus iii) is proved. Finally if, for a continuum $c \subset Q$, we have $cQ^* \neq 0$, $\operatorname{diam} x(c) < D$, then for any w, $\bar{w}$, $w \in c$, $\bar{w} \in cQ^*$, we have

$$\begin{aligned} |x(w) - x_0| &\leq |x(w) - x(\bar{w})| + |x(\bar{w}) - x(w_0)| \\ &\leq \operatorname{diam} x(c) + \operatorname{diam} C < D + D = 2D; \end{aligned}$$

hence $x(c) \subset B(2D)$ and $c \subset Q - \pi$. Thus also v) is proved. Lemma 19 is now completely proved.

§ 3. Existence Theorems.

25. THEOREM I. *Let A be a closed bounded convex set of E_3 and C a closed Jordan curve in A. Let W be the family of all oriented Fréchet surfaces S in A whose Lebesgue area is finite and whose boundary curve is C,*

and let us suppose W not empty. Then every positive definite and semi-regular integral $\mathfrak{J}(S)$ has an absolute minimum in W.

The proof follows in nos. 26-41.

26. The function $F(x, p)$ is continuous in the closed bounded set I of all points (x, p) such that $x \in A$, $|p| = 1$, and $F > 0$ for any $(x, p) \in I$. Then if m, M are the minimum and the maximum of F in I, we have $0 < m \leq F(x, p) \leq M$ for any $(x, p) \in I$ and therefore $m|p| \leq F(x, p) \leq M|p|$ for any $x \in A$ and any $|p| \neq 0$. By a theorem of McShane [18] we can define for any $x \in E_3$ and $|p| = 1$ a function $F_0(x, p)$ continuous at each point (x, p) and coinciding with $F(x, p)$ for any $(x, p) \in I$. In addition we can suppose $m \leq F_0(x, p) \leq M$ for any $x \in E_3$, $|p| = 1$. Let us define $F_0(x, p)$ for each $|p| \neq 0$ by putting $F_0(x, p) = |p| F_0(x, p/|p|)$. We have $F_0(x, p) = F(x, p)$ for any $x \in A$, $|p| \neq 0$. Indeed, by no. 7, b, $F_0(x, p) = |p| F_0(x, p/|p|) = |p| F(x, p/|p|) = F(x, p)$. The function $F_0(x, p)$ satisfies all the conditions of no. 7. By $\mathfrak{J}(S)$ we mean now $\mathfrak{J}(S) = \int\int_S F_0(x, p)\,dw$ and, for any surface S such that $[S] \subset A$, we have $\mathfrak{J}(S) = \int\int_S F(x, p)\,dw$.

We observe that $m \leq F_0(x, p) \leq M$ for any $x \in E_3$, $|p| = 1$, and $m|p| \leq F_0(x, p) \leq M|p|$ for any $x \in E_3$, $|p| \neq 0$. In consequence, by [2], we have also

$$mL(S) \leq \mathfrak{J}(S) \leq ML(S) \tag{1}$$

for any surface S.

27. For any $\delta \geq 0$ let A_δ be the closed convex bounded set of all points $x \in E_3$ such that $\{x, A\} \leq \delta$. We have $A_\delta \supset A$ and $A_0 = A$. For any $\delta \geq 0$ let $W(\delta)$ be the family of all surfaces S such that $[S] \subset A_\delta$, $\| \partial S, C \| \leq \delta$. We have $W(\delta) \supset W$, $W(0) = W$, and $W(\delta') \supset W(\delta)$ for any $0 \leq \delta \leq \delta'$. W is not empty, hence also $W(\delta)$ is not empty for any $\delta > 0$. Let us denote by $i, j(\delta)$ the g. l. b. of $\mathfrak{J}(S)$ when $S \in W$, or $S \in W(\delta)$, $\delta > 0$. Then $0 \leq j(\delta) \leq i < +\infty$ and $j(\delta') \leq j(\delta)$ for any $0 < \delta \leq \delta'$. Hence there exists $j = \lim j(\delta)$ as $\delta \to 0 + 0$ and we have $0 \leq j(\delta) \leq j \leq i$ for any $\delta > 0$. For any integer $n = 1, 2, \cdots$, let

$$\epsilon_n = 2^{-n}, \qquad \mu_n = \min[(n+2)^{-1}2^{-n-2}, 2^{-2}M\epsilon_n^2],$$

and let $\delta_n > 0$ be a real number such that $j(5\delta_n) > j - \mu_n$, $\delta_n \leq 1/n$.

Finally let $\alpha_n' = \alpha\,(C, \delta_n)$ be the number of Lemma 11 related to the Jordan curve C and the number $\epsilon = \delta_n$. Let $\alpha_n = \min[\alpha_n', \delta_n]$.

28. Let us consider the family $W(\alpha_n)$. There is a surface $S_n' \in W(\alpha_n)$ such that

$$j(\alpha_n) \leq \mathfrak{J}(S_n') \leq j(\alpha_n) + \mu_n,$$

and we have $[S_n'] \subset A_{\alpha_n}$, $\| \partial S_n', C \| \leq \alpha_n$. By (1) we have $L(S_n') \leq (1/m)\mathfrak{J}(S_n')$. By Lemma 10 the surface S_n' can be retracted to a base surface S_n'' having the same boundary $\partial S_n'' = \partial S_n'$ and such that $[S_n''] \subset [S_n']$, $L(S_n'') \leq L(S_n')$, $\mathfrak{J}(S_n'') \leq \mathfrak{J}(S_n')$. Hence $\| \partial S_n'', C \| = \| \partial S_n', C \| \leq \alpha_n$. By Lemma 11 the surface S_n'' can be retracted to a non-degenerate surface S_n''' whose boundary $\partial S_n'''$ is a continuous curve such that $\| \partial S_n''', C \| \leq \delta_n$. In addition

$$[S_n'''] \subset [S_n''], \qquad L(S_n''') \leq L(S_n''), \qquad \mathfrak{J}(S_n''') \leq \mathfrak{J}(S_n'').$$

We observe that $\delta_n \leq 1/n$, hence $\| \partial S_n''', C \| \to 0$ as $n \to \infty$. On the other hand the surfaces S_n''', $n = 1, 2, \cdots$, are non-degenerate surfaces whose areas $L(S_n''')$ satisfy the relation

$$\begin{aligned} L(S_n''') \leq L(S_n'') \leq L(S_n') &\leq (1/m)\mathfrak{J}(S_n') \\ &\leq (1/m)[j(\alpha_n) + \mu_n] \leq (1/m)(j+1) < +\infty. \end{aligned}$$

By Lemma 12 there are generalized conformal D-representations S_n''': $x = x_n'''(w)$, $w \in Q$, of each S_n''' on Q, satisfying a three points condition on Q^* and (no. 5) $D[x_n'''] = L(S_n''')$, $n = 1, 2, \cdots$. By Lemma 6 and Lemma 8 there is, for each n, a polyhedral surface S_n and a quasi-linear representation $S_n : x = x_n(w)$, $w \in Q$, of S_n on Q, such that $| x_n(w) - x_n'''(w) | \leq \delta_n$ for any $w \in Q$, $D[x_n(w)] < D[x_n'''(w)] + \mu_n$, $| \mathfrak{J}(S_n) - \mathfrak{J}(S_n''') | < \mu_n$.

In the first place we have $| x_n(w) - x_n'''(w) | \leq \delta_n$ for any $w \in Q^*$; hence also the vectors $x_n(w)$, $n = 1, 2, \cdots$, satisfy a three points condition on Q^*. In addition, if $C_n = \partial S_n$ is the boundary curve of S_n, we have $\| C_n, \partial S_n''' \| = \| \partial S_n, \partial S_n''' \| \leq \delta_n$ and $\| C_n, C \| \leq \| C_n, \partial S_n''' \| + \| \partial S_n''', C \| \leq \delta_n + \delta_n = 2\delta$. Finally we have

$$[S_n] \subset [S_n''']_{\delta_n} \subset [S_n'']_{\delta_n} \subset [S_n']_{\delta_n} \subset [A_{\alpha_n}]_{\delta_n} \subset A_{\alpha_n + \delta_n} \subset A_{2\delta_n}.$$

This implies that $S_n \in W(2\delta_n)$, hence $\mathfrak{J}(S_n) \geq j(2\delta_n) \geq j - \mu_n$. By all the preceding considerations we have

$$\begin{aligned} \mathfrak{J}(S_n) \leq \mathfrak{J}(S_n''') + \mu_n \leq \mathfrak{J}(S_n'') + \mu_n &\leq \mathfrak{J}(S_n') + \mu_n \\ &\leq j(\alpha_n) + 2\mu_n \leq j + 2\mu_n, \end{aligned}$$

and, by (1) and no. 5, also

$$a(S_n) \leq D[x_n(w)] \leq D[x_n'''(w)] + \mu_n = L(S_n''') + \mu_n$$
$$< (1/m)(j+1) + 1.$$

We can assemble the properties of the polyhedral surfaces S_n and of their quasi-linear representations $S_n: x = x_n(w)$, $w \in Q$, (satisfying a three points condition on Q^*) as follows:

$$\| C_n, C \| \leq 2\delta_n, \qquad [S_n] \subset A_{2\delta_n}, \qquad S_n \in W(2\delta_n),$$
$$a(S_n) \leq D[x_n] < (1/m)(j+1) + 1, \qquad j - \mu_n \leq \mathfrak{J}(S_n) \leq j + 2\mu_n,$$
$$n = 1, 2, \cdots.$$

29. The representations $x = x_n(w)$, $w \in Q$, of S_n on Q are quasi-linear, hence they are D-representations and, by no. 28, $D[x_n(w)] < (1/m)(j+1) + 1$. The constant $N = (1/m)(j+1) + 1$ does not depend upon n. C is a Jordan curve and $\| C_n, C \| \to 0$ as $n \to \infty$, hence, by Lemma 14, the vectors $x_n(w)$, $w \in Q$, are equicontinuous on Q^*. There is a subsequence $x_{n_m}(w)$, $w \in Q$, of vectors $x_n(w)$ which converges uniformly on Q^*. Let us suppose, for simplicity's sake, that $n_m = m$. Let $X(w) = \lim x_n(w)$, $w \in Q^*$, as $n \to \infty$. If $C_0: x = X(w)$, $w \in Q^*$, we have $\| C_n, C_0 \| \to 0$, $\| C_n, C \| \to 0$, $\| C, C_0 \| \leq \| C, C_n \| + \| C_n, C_0 \|$, hence $\| C, C_0 \| = 0$. That is, $x = X(w)$, $w \in Q^*$, gives a representation of the Jordan curve C. By the equicontinuity of the vectors $x_n(w)$, $w \in Q^*$, $n = 1, 2, \cdots$, we can determine, for each integer ν ($\nu = 1, 2, \cdots$), a real number $\tau_\nu > 0$ such that

$$| x_n(w) - x_n(w') | \leq \min[\epsilon_\nu, \delta_\nu], \qquad n = 1, 2, \cdots,$$

for any $w, w' \in Q^*$, $| w - w' | \leq \tau_\nu$. The numbers τ_ν depend upon ν but not on n.

30. Let us consider again the D-vectors $x_n(w)$, $w \in Q$, $n = 1, 2, \cdots$, which satisfy the relation $D[x_n] < N$, N constant, $n = 1, 2, \cdots$, and let us apply Lemma 13. Let ν indicate any integer, $\nu = 1, 2, \cdots$, and $\xi_\nu > 0$ be any real number. Then, by Lemma 13, with $\epsilon = \xi_\nu$, we can determine, for each ν, real numbers $\eta_\nu > 0$, $\delta_{n\nu} > 0$, $\eta_\nu < \delta_{n\nu} < \xi_\nu$, and regular subdivisions (no. 4) $\Delta_{n\nu}$ of Q in rectangles r whose dimensions lie between $\delta_{n\nu}$ and $2\delta_{n\nu}$, such that the image of each side of the rectangles $r \in \Delta_{n\nu}$ *not* on Q^*, is a rectifiable curve of length $\leq \xi_\nu$. We observe that the numbers η_ν depend only on ν, and $\delta_{n\nu}$ on ν and n.

As the application of Lemma 13 can be made successively for $\nu = 1, 2, \cdots$, we can choose

$$\xi_\nu = \min[\tau_\nu/2, \delta_\nu, \epsilon_\nu, \eta_{\nu-1}/2^3, \eta_{\nu-2}/2^4, \cdots, \eta_1/2^{\nu+1}], \qquad \nu = 1, 2, \cdots .$$

For each ν the dimensions of all rectangles $r \in \Delta_{n\nu}$, $n = 1, 2, \cdots$, lie between $\delta_{n\nu}$ and $2\delta_{n\nu}$, hence between η_ν and $2\xi_\nu < \eta_{\nu-1}$. Therefore they are less than the dimensions of all rectangles $r \in \Delta_{n,\nu-1}$, $n = 1, 2, \cdots$. The images of the sides on Q^* of the rectangles $r \in \Delta_{n\nu}$ under the vector $x_n(w)$ are arcs of diameter $\leq \min[\epsilon_\nu, \delta_\nu]$. The images of all other sides of the rectangles $r \in \Delta_{n\nu}$ under the vector $x_n(w)$ are arcs of length $\leq \min[\epsilon_\nu, \delta_\nu]$. The image of the boundary r^* of each $r \in \Delta_{n\nu}$ is a curve of diameter $\leq \min[4\epsilon_\nu, 4\delta_\nu]$ if $r^*Q^* \neq 0$, and is a curve of length $\leq \min[4\epsilon_\nu, 4\delta_\nu]$ if $r^*Q^* = 0$. In any case we have $\operatorname{diam} x_n(r^*) \leq 4\epsilon_\nu$ for any $r \in \Delta_{n\nu}$, $\nu = 1, 2, \cdots$, $n = 1, 2, \cdots$.

31. In the following we suppose n an arbitrary integer ($n = 1, 2, \cdots$), and $\nu = 1, 2, \cdots, n$. Let us indicate (no. 4) by $w_{ij} \equiv (u_i, v_j)$, $i = 0, 1, 2, \cdots, \alpha + 1$, $j = 0, 1, 2, \cdots, \beta + 1$, $u_0 = v_0 = 0$, $u_{\alpha+1} = v_{\beta+1} = 1$, all vertices of the rectangles $r \in \Delta_{n\nu}$. Let us denote by τ each segment ($u = u_i$, $v_j \leq v \leq v_{j+1}$), $i = 1, \alpha$, $j = 1, 2, \cdots, \beta - 1$, or ($u_i \leq u \leq u_{i+1}$, $v = v_j$), $i = 1, 2, \cdots, \alpha - 1$, $j = 1, \beta$. Let τ' be all analogous segments $\tau' \subset Q^*$ with $i = 0, \alpha + 1$, $j = 0, 1, \cdots, \beta$, or $i = 0, 1, \cdots, \alpha$ $j = 0, \beta + 1$. The segments τ constitute the boundary $R_{n\nu}{}^*$ of the rectangle $R_{n\nu} \equiv [u_1 \leq u \leq u_\alpha, v_1 \leq v \leq v_\beta]$, the segments τ' constitute the boundary Q^* of Q. Let us indicate by t and t' the polygonal lines which are images of the segments τ and τ' under the vector $x_n(w)$. Let $S_{n\nu}{}'$ be the polyhedral surface $S_{n\nu}{}' : x = x_n(w)$, $w \in R_{n\nu}$. Then the sums of the arcs t, or t', constitute the boundary curves $\partial S_{n\nu}{}'$, $\partial S_n = C$ of the surfaces $S_{n\nu}{}'$, S_n. Clearly $\mathfrak{S}(S_{n\nu}{}') \leq \mathfrak{S}(S_n)$. Let us make correspond to each segment $\tau \subset R_{n\nu}{}^*$ with $i = 1, \alpha$, $j = 1, 2, \cdots, \beta - 1$, the segment $\tau' \subset Q^*$ with $i = 0, \alpha + 1$ and the same j; let us make correspond to each segment $\tau \subset R_{n\nu}{}^*$ with $i = 1, 2, \cdots, \alpha - 1$, $j = 1, \beta$, the segment $\tau' \subset Q^*$ with $j = 0, \beta + 1$ and the same i (the correspondences being linear); let us make correspond to the vertices $w_{11}, w_{\alpha 1}, w_{\alpha\beta}, w_{1\beta}$ of $R_{n\nu}{}^*$ the arcs of Q^* which are sums of the two segments $\tau' \subset Q^*$ adjacent to the vertices $w_{00}, w_{\alpha+1,0}, w_{\alpha+1,\beta+1}$, $w_{0,\beta+1}$ of Q^*. We have a continuous monotone correspondence between $R_{n\nu}{}^*$ and Q^*. If $w \in R_{n\nu}{}^*$, $w' \in Q^*$ are corresponding points, we have

$$\begin{aligned} |x_n(w) - x_n(w')| &\leq |x_n(w) - x_n(\bar{w})| \\ &\quad + |x_n(\bar{w}) - x_n(\bar{w}')| + |x_n(\bar{w}') - x_n(w')|, \end{aligned}$$

where, if $w \in \tau$, $w' \in \tau'$, then $\bar{w}$, $\bar{w}'$ are corresponding end points of τ and τ';

analogously if w is a vertex of $R_{n\nu}{}^*$ and $w' \varepsilon \tau'$, $\bar{w} = w$. By nos. 29, 30, we have $|x_n(w) - x_n(w')| \leq \xi_\nu + \xi_\nu + \delta_\nu \leq 3\delta_\nu$. This assures us that $\| \partial S_{n\nu}', \partial S_n \| \leq 3\delta_\nu$ and therefore $\| \partial S_{n\nu}', C \| \leq \| \partial S_{n\nu}', \partial S_n \| + \| \partial S_n, C \| \leq 3\delta_\nu + 2\delta_\nu = 5\delta_\nu$. On the other hand $[S_{n\nu}'] \subset [S_n] \subset A_{2\delta_n} \subset A_{2\delta_\nu}$, and finally, by no. 27, $S_{n\nu}' \varepsilon W(5\delta_N)$ and

$$j - \mu_\nu \leq \mathfrak{J}(S_{n\nu}') \leq \mathfrak{J}(S_n) \leq j + 2\mu_n \leq j + 2\mu_\nu.$$

In addition,

$$0 \leq \mathfrak{J}(S_n) - \mathfrak{J}(S_{n\nu}') \leq 3\mu_\nu.$$

32. For each rectangle $r \varepsilon \Delta_{n\nu}$ let us call Σ the polyhedral surface $\Sigma: x = x_n(w)$, $w \varepsilon r$, and $L = \partial\Sigma$. Let L denote also the length of $\partial\Sigma$. Let us prove that $a(\Sigma) \leq (5M/m)\epsilon_\nu{}^2$. First let us suppose $r \subset R_{n\nu}$; thus $L \leq 4\epsilon_\nu$. If the above statement were false then, for a certain $r \varepsilon \Delta_{n\nu}$, we would have $a(\Sigma) > (5M/m)\epsilon_\nu{}^2$. Thus, by Lemma 16 applied to the curve $L: x = x_n(w)$, $w \varepsilon r^*$, there is a polyhedral surface Σ^* contained in the minimum convex body containing L, of area $a(\Sigma^*) \leq (1/4)L^2 \leq (1/4)(4\epsilon_\nu)^2 = 4\epsilon_\nu{}^2$ and a quasi-linear representation $\Sigma^*: x = x^*(w)$, $w \varepsilon r$, of Σ^* on r with $x^*(w) = x_n(w)$, $w \varepsilon r^*$. If we denote by $S_{n\nu}{}^*: x = x_{n\nu}{}^*(w)$, $w \varepsilon R_{n\nu}$, the surfaces that we obtain by putting $x_{n\nu}{}^*(w) = x_n(w)$, $w \varepsilon R_{n\nu} - r$, $x_{n\nu}{}^*(w) = x^*(w)$, $w \varepsilon r$, then we have $\partial S_{n\nu}{}^* = \partial S_n$ and $[S_{n\nu}{}^*]$ is contained in the same convex body $A_{2\delta_n}$ in which $S_{n\nu}$ is already contained. Hence $\| \partial S_{n\nu}{}^*, C \| = \| \partial S_{n\nu}', C \| \leq 5\delta_\nu$, $[S_{n\nu}{}^*] \subset A_{2\delta_\nu}$ and consequently $S_{n\nu}{}^* \subset W(5\delta_\nu)$ and $\mathfrak{J}(S_{n\nu}{}^*) \geq j - \mu_\nu$. Now, by no. 27, we have

$$\begin{aligned} j - \mu_\nu \leq \mathfrak{J}(S_{n\nu}{}^*) &= \mathfrak{J}(S_{n\nu}{}^*) + \mathfrak{J}(\Sigma^*) - \mathfrak{J}(\Sigma) \leq (j + 2\mu_\nu) + Ma(\Sigma^*) - ma(\Sigma) \\ &< j + 2\mu_\nu + 4M\epsilon_\nu{}^2 - m(5M/m)\epsilon_\nu{}^2 = j + 2\mu_\nu - M\epsilon_\nu{}^2 \\ &\leq j + 2\mu_\nu - 4\mu_\nu = j - 2\mu_\nu. \end{aligned}$$

which is contradictory. This proves $a(\Sigma) \leq (5M/m)\epsilon_\nu{}^2$ for any $r \varepsilon \Delta_{n\nu}$, $r \subset R_{n\nu}$, $\nu = 1, 2, \cdots, n$, $n = 1, 2, \cdots$. Let us now suppose $r \subset Q - R_{n\nu}$; then $a(\Sigma) \leq (1/m)\mathfrak{J}(\Sigma) = (1/m)[\mathfrak{J}(S_n) - \mathfrak{J}(S_{n\nu}')] \leq 3\mu_\nu/m$, and, by no. 27, $a(\Sigma) \leq (3/m)(2^{-2}M\epsilon_\nu{}^2) < (5M/m)\epsilon_\nu{}^2$ for any $r \varepsilon \Delta_{n\nu}$, $r \subset Q - R_{n\nu}$, $\nu = 1, 2, \cdots, n$, $n = 1, 2, \cdots$. The statement at the beginning of this no. is thus proved.

33. For each $n = 1, 2, \cdots$ and $\nu = 1, 2, \cdots, n$, let us apply Lemma 19 in connection with each rectangle $r \varepsilon \Delta_{n\nu}$, the vector $x_n(w)$, $w \varepsilon Q$, the constant $K = M/m \geq 1$, any point $x_0 = x_n(w_0)$, where $w_0 \varepsilon r^*$, and $D = 4\epsilon_\nu$.

This is possible because from no. 30 we have diam $x_n(r^*) \leq 4\epsilon_\nu = D$ for any $r \in \Delta_{n\nu}$. Let $T_{n\nu}$ denote the collection of simple polygons $\pi \subset r$ defined by Lemma 19 in all rectangles $r \in \Delta_{n\nu}$. According to Lemma 19 all polygons π are interior to r together with their boundaries. Let $x_{n\nu}(w)$ be the vector defined by Lemma 19 on each rectangle $r \in \Delta_{n\nu}$. Because of $x_{n\nu}(w) = x_n(w)$ for any $w \in r^*$, as well as for any $w \in r - \Sigma\pi$, the vectors $x_{n\nu}(w)$, $w \in r$, $r \in \Delta_{n\nu}$ define a unique vector $x_{n\nu}(w)$, $w \in Q$, continuous and quasi-linear on Q. Let $S_{n\nu}$ be the polyhedral surface $S_{n\nu}: x = x_{n\nu}(w)$, $w \in Q$.

For each rectangle $r \in \Delta_{n\nu}$ let σ, σ_0 be the family of polyhedral surfaces defined by the vectors $x_n(w)$ and $x_{n\nu}(w)$ on the family of polygons π, $\pi \in T_{n\nu}$, $\pi \subset r$. We have $\partial\sigma = \partial\sigma_0$. Let $\mathfrak{L}$ be the total length of $\partial\sigma = \partial\sigma_0$, i. e., the sum of the lengths p of the common boundaries curves of the polyhedral surfaces defined by $x_n(w)$, $x_{n\nu}(w)$ on the polygons π, $\pi \in T_{n\nu}$. $\pi \subset r$, $r \in \Delta_{n\nu}$. We have $\mathfrak{L} = \Sigma p$, where the sum ranges over the polygons $\pi \in T_{n\nu}$, $\pi \subset r$.

34. Since $x_{n\nu}(w) = x_n(w)$ for any $w \in Q^*$ we have $\partial S_{n\nu} = \partial S_n$, $\| \partial S_{n\nu}, C \| = \| \partial S_n, C \| = \| C_n, C \| \leq 2\delta_n$. Each surface σ_0 is by Lemma 19 contained in the minimum convex body containing $[\mathfrak{L}] = [\partial\sigma] = [\partial\sigma_0]$ and, because of $[\mathfrak{L}] \subset [S_n] \subset A_{2\delta_n}$ where $A_{2\delta_n}$ is a convex body, we have $[S_{n\nu}] \subset A_{2\delta_n}$, $\nu = 1, 2, \cdots, n$. This assures us (no. 27) that $S_{n\nu} \in W_{2\delta_n}$, hence $\mathfrak{J}(S_{n\nu}) \geq j(2\delta_n) \geq j - \mu_n$, $\nu = 1, 2, \cdots, n$.

By Lemma 19 we have

$$a(\sigma_0) \leq (K^{-1}/4)a(\sigma) = (m/4M)a(\sigma).$$

By (1) (no. 26) it follows that

$$0 \leq \mathfrak{J}(\sigma_0) \leq Ma(\sigma_0) \leq M(m/4M)a(\sigma) \leq M(m/4M)(1/m)\mathfrak{J}(\sigma),$$

hence

$$0 \leq \mathfrak{J}(\sigma_0) \leq (1/4)\mathfrak{J}(\sigma). \tag{2}$$

Let us indicate by $S_{n\nu 0}$ the polyhedral surface (sum of triangles) defined by the vectors $x_{n\nu}(w)$, $x_n(w)$ on the polygonal region $Q - \Sigma\pi$, $\pi \in T_{n\nu}$. We have $S_{n\nu 0} + \Sigma\sigma = S_n$, $S_{n\nu 0} + \Sigma\sigma_0 = S_{n\nu}$, where the sum ranges on all families of surfaces σ [resp. σ_0], each family determined by the vector x_n [resp. $x_{n\nu}$] on the polygons π, $\pi \subset T_{n\nu}$, $\pi \subset r$, for any $r \in \Delta_{n\nu}$. We have

$$\mathfrak{J}(S_{n\nu 0}) + \sum_{r \in \Delta_{n\nu}} \mathfrak{J}(\sigma) = \mathfrak{J}(S_n), \qquad \mathfrak{J}(S_{n\nu 0}) + \sum_{r \in \Delta_{n\nu}} \mathfrak{J}(\sigma_0) = \mathfrak{J}(S_{n\nu}); \tag{3}$$

and, by (2), we have $\mathfrak{J}(S_{n\nu}) \leq \mathfrak{J}(S_n)$; hence $\mathfrak{J}(S_{n\nu}) \leq \mathfrak{J}(S_n) \leq j + 2\mu_n$, $\nu = 1, 2, \cdots, n$. We have thus proved that

$$j - \mu_n \leq \mathfrak{J}(S_{n\nu}) \leq j + 2\mu_n, \qquad \nu = 1, 2, \cdots, n.$$

By (2) and (3) we have also

$$3\mu_n \geq \mathfrak{F}(S_n) - \mathfrak{F}(S_{n\nu}) = \sum_{r \in \Delta_{n\nu}} [\mathfrak{F}(\sigma) - \mathfrak{F}(\sigma_0)]$$

$$\geq \sum_{r \in \Delta_{n\nu}} (1 - 1/4)\mathfrak{F}(\sigma) = (3/4) \sum_{r \in \Delta_{n\nu}} \mathfrak{F}(\sigma),$$

hence

$$\sum_{r \in \Delta_{n\nu}} \mathfrak{F}(\sigma) \leq 4\mu_n, \qquad \nu = 1, 2, \cdots, n. \tag{4}$$

If we denote by $\sigma_{n\nu}$, $\sigma_{n\nu 0}$ the total families of all surfaces σ, σ_0 for all $r \in \Delta_{n\nu}$, we have also, by (2) and (4),

$$\mathfrak{F}(\sigma_{n\nu 0}) \leq (1/4)\mathfrak{F}(\sigma_{n\nu}) \leq \mu_n, \qquad \nu = 1, 2, \cdots, n.$$

35. By (1) and Lemma 19 we have

$$a(\sigma) \leq (1/m)\mathfrak{F}(\sigma), \qquad \mathfrak{L}^2 \leq (1/K)a(\sigma) = (m/M)a(\sigma), \tag{5}$$

where $\mathfrak{L} = \partial\sigma = \partial\sigma_0$ and we have denoted by $\mathfrak{L}$ also the length of $\mathfrak{L}$. From (4) and (5) we have

$$\sum_{r \in \Delta_{n\nu}} \mathfrak{L}^2 \leq (m/M) \sum_{r \in \Delta_{n\nu}} a(\sigma) \leq (1/M) \sum_{r \in \Delta_{n\nu}} \mathfrak{F}(\sigma) \leq (1/M)4\mu_n,$$

hence

$$\sum_{r \in \Delta_{n\nu}} \mathfrak{L}^2 \leq 4\mu_n/M, \qquad \sum_{r \in \Delta_{n\nu}} a(\sigma) \leq 4\mu_n/m;$$

and also, by no. 27,

$$\sum_{r \in \Delta_{n\nu}} \Big(\sum_{\substack{\pi \in T_{n\nu} \\ \pi \subset r}} p \Big)^2 = \sum_{r \in \Delta_{n\nu}} \mathfrak{L}^2 \leq 4\mu_n/\mathrm{M} \leq \epsilon_n^2 < \epsilon_n, \qquad \nu = 1, 2, \cdots, n.$$

In particular, for each rectangle $r \in \Delta_{n\nu}$ the total length $\mathfrak{L}$ of the common boundary $\mathfrak{L} = \partial\sigma = \partial\sigma_0$ of the surfaces defined by the vectors $x_n(w)$, $x_{n\nu}(w)$ on the polygons π, $\pi \in T_{n\nu}$, $\pi \subset r$, satisfies the relation

$$\mathfrak{L} \leq [4\mu_n/M]^{1/2} \leq \epsilon_n, \qquad \nu = 1, 2, \cdots, n.$$

For any $\nu = 1, 2, \cdots, n$, we have $\nu \leq n$, $\epsilon_\nu \geq \epsilon_n$, hence $\mathfrak{L} \leq \epsilon_\nu$. For any polygon $\pi \in T_{n\nu}$ we have $\pi \subset r$, $r \in \Delta_{n\nu}$ and the diameter of the set $x_n(\pi^*)$ is $\leq$ the length of the curve $x = x_n(w)$, $w \in \pi^*$. But this curve is the boundary curve of one of the surfaces which are part of σ. Hence diam $x_n(\pi^*) \leq$ length $(\partial\sigma) = \mathfrak{L} \leq \epsilon_n \leq \epsilon_\nu$, $\nu = 1, 2, \cdots, n$.

36. For any $r \in \Delta_{n\nu}$ let us indicate by Σ, Σ_0 the polyhedral surfaces

$\Sigma: x = x_n(w)$, $w \in r$, $\Sigma_0: x = x_{n\nu}(w)$, $w \in r$. Then Σ_0 (Lemma 19) is completely contained in a sphere of center $x_0 = x_n(w_0)$, $w_0 \in r^*$, and radius $\rho \leq 2D + 3[Ka(\Sigma)]^{1/2}$ whence, by no. 32, $\rho \leq 8\epsilon_\nu + 3[(M/m)(5M/m)\epsilon_\nu^2]^{1/2} < 15K\epsilon_\nu$. Therefore $\operatorname{diam} x_{n\nu}(r) = \operatorname{diam}(\Sigma_0) < 30K\epsilon_\nu$ for any $r \in \Delta_{n\nu}$, $\nu = 1, 2, \cdots, n$, $n = 1, 2, \cdots$, $K = M/m \geq 1$.

37. Let $T'_{n1} \subset T_{n1}$ be the subcollection of all polygons $\pi \in T_{n1}$ which are not completely contained in any rectangle $r \in \Delta_{n2} + \Delta_{n3} + \cdots + \Delta_{nn}$; let $T'_{n\nu} \subset T_{n\nu}$, $\nu = 2, 3, \cdots, n-1$, be the subcollection of all polygons $\pi \in T_{n\nu}$ which are not completely contained in any rectangle $r \in \Delta_{n,\nu+1} + \Delta_{n,\nu+2} + \cdots + \Delta_{nn}$ and not completely contained in the set $T'_{n1} + T'_{n2} + \cdots + T'_{n,\nu-1}$; let $T'_{nn} \subset T_{nn}$ be the subcollection of all polygons $\pi \in T_{nn}$ which are not completely contained in $T'_{n1} + T'_{n2} + \cdots + T'_{n,n-1}$.

Let us denote by $T_{n\nu}$, $T'_{n\nu}$ also the (closed) sets of points of Q covered by the polygons $\pi \in T_{n\nu}$, $T'_{n\nu}$. Let us denote by $\Lambda_{n\nu} \subset \Delta_{n\nu}$ the subcollection of all rectangles $r \in \Delta_{n\nu}$ which are not completely contained in the closed set $T'_{n1} + T'_{n2} + \cdots + T'_{n,\nu-1}$, $\nu = 1, 2, \cdots, n$. For any polygon $\pi \in T'_{n\nu}$ we have $\pi \subset r$, $r \in \Delta_{n\nu}$. We prove the following statement: π *is completely exterior to the set* $r(T'_{n1} + T'_{n2} + \cdots + T'_{n,\nu-1})$. First let us observe that $\pi \in T'_{n\nu}$, hence π is not completely contained in $T'_{n1} + T'_{n2} + \cdots + T'_{n,\nu-1}$. Here $\pi \subset r$, hence also r is not completely contained in $T'_{n1} + T'_{n2} + \cdots + T'_{n,\nu-1}$. Let π' be any polygon $T'_{n1} + \cdots + T'_{n,\nu-1}$; then r is not completely contained in π'. On the other hand, by definition of $T'_{n1}, T'_{n2}, \cdots, T'_{n,\nu-1}$, π' is not completely contained in r, because $r \in \Delta_{n\nu}$. If r and π' have no interior points in common the statement is trivial; if r and π' have interior points in common then they must have also points of their boundaries in common, that is $\pi'^* r^* \neq 0$ and each component of $\pi'^* r$ is a polygonal line having both end points on r^*. From no. 35 we have $\operatorname{diam} x_n(\pi'^*) \leq \epsilon_n \leq \epsilon_\nu < D$. By Lemma 19, statement v) we deduce that π'^* is completely outside of π. Because π is in r but not completely contained in π', it follows that π is outside of π'. This holds for all polygons $\pi' \in T'_{n1} + T'_{n2} + \cdots + T'_{n,\nu-1}$; hence π is outside of $r(T'_{n1} + T'_{n2} + \cdots + T'_{n,\nu-1})$ and the statement is proved.

38. By no. 37 no polygon $\pi \in T_{n\nu}'$ has points in common with polygons $\pi' \in T_{n1}' + \cdots + T_{n,\nu-1}'$, $\nu = 2, 3, \cdots, n$. Let us denote by $X_n(w)$, $w \in Q$, the vector which is $= x_{n\nu}(w)$ for any $w \in \pi$, $\pi \in T_{n\nu}'$, $\nu = 1, 2, \cdots, n$, and which is $= x_n(w)$ for any $w \in Q - I_n'$, $I_n' = T_{n1}' + \cdots + T_{nn}'$. The vector $X_n(w)$, $w \in Q$, is continuous and quasi-linear on Q. Let S_n be the polyhedral

surface $\bar{S}_n : x = X_n(w)$, $w \in Q$. The same reasoning as at the beginning of no. 34 assures us that $\partial \bar{S}_n = \partial S_n = C_n$, $\| \partial \bar{S}_n, C \| \leq 2\delta_n$, $[\bar{S}_n] \subset A_{2\delta_n}$, hence $\bar{S}_n \in W(2\delta_n)$ and $\mathfrak{J}(\bar{S}_n) \geq j - \mu_n$. The vector $X_n(w)$ defines on the polygons $\pi \in T_{n\nu}'$ a family $\sigma_{n\nu 0}'$ of polyhedral surfaces which, because of $T_{n\nu}' \subset T_{n\nu}$, is only a part of the family $\sigma_{n\nu 0}$ (no. 34), defined by $x_{n\nu}(w)$ on *all* polygons $\pi \in T_{n\nu}$ and which is the sum of the surfaces σ_0 of nos. 33-34 for all $r \in \Delta_{n\nu}$. If we denote by $\bar{S}_{n0}$ the surface (sum of triangles) defined by $X_n(w)$, as well as by $x_n(w)$, on $Q - I_n'$, then, by no. 28 and no. 34, we have

$$\mathfrak{J}(\bar{S}_n) = \mathfrak{J}(\bar{S}_{n0}) + \sum_{\nu=1}^{n} \mathfrak{J}(\sigma_{n\nu 0}') \leq \mathfrak{J}(S_n) + \sum_{\nu=1}^{n} \mathfrak{J}(\sigma_{n\nu 0})$$

$$\leq j + 2\mu_n + \sum_{\nu=1}^{n} \mu_n = j + (n+2)\mu_n.$$

By no. 27 we have finally

$$j - 2^{-n} < j - \mu_n \leq \mathfrak{J}(\bar{S}_n) \leq j + (n+2)\mu_n < j + 2^{-n}.$$

39. Let $\epsilon > 0$ be any real number. Let $\mu = \mu(\epsilon)$ be the first integer such that $132K\epsilon_\mu < \epsilon$. Let $\delta = \delta(\epsilon) = \eta_\mu > 0$. Let w_1, w_2 be any two points of Q with $| w_1 - w_2 | < \delta$ and let $n \geq \mu$ be any integer. Then $w_1 \in r_1$, $w_2 \in r_2$, $r_1, r_2 \in \Delta_{n\mu}$, where r_1, r_2 are the same rectangle of $\Delta_{n\mu}$, or two rectangles of $\Delta_{n\mu}$ having one side or one vertex in common. In any case $r_1^* r_2^* \neq 0$. We are going to consider several cases:

a) w_1, w_2 belong to the same polygon $\pi \in T_{n\nu}'$, $1 \leq \nu \leq n$. Then, by no. 38, $| X_n(w_1) - X_n(w_2) | = | x_{n\nu}(w_1) - x_{n\nu}(w_2) |$ and the last expression is $\leq$ the diameter of the minimum convex body containing $x_n(\pi^*)$, which is certainly ≤ 2 diam $x_n(\pi^*)$. Finally, by no. 35, we have

$$| X_n(w_1) - X_n(w_2) | < 2 \text{ diam } x_n(\pi^*) \leq 2\epsilon_n \leq 2\epsilon_\mu < \epsilon.$$

b) $w_i \in \pi_i$, $\pi_i \in T'_{n\nu_i}$ π_i not completely contained in r_i, $r_i \in \Delta_{n\mu}$, $i = 1, 2$, $\pi_1 \neq \pi_2$. Here $(r_1^* + r_2^*)\pi_i^*$ is not empty; hence let w_i' be any point $w_i' \in (r_1^* + r_2^*)\pi_i^*$, $i = 1, 2$. We have $| X_n(w_i) - X_n(w_i') | \leq 2\epsilon_n$, $i = 1, 2$. The two points w_1', w_2' are in $r_1^* + r_2^*$; hence there is a polygonal line λ joining w_1', w_2' contained in $r_1^* + r_2^*$, and therefore formed by at most four consecutive segments contained each in one side of the rectangles r_1, r_2. By no. 30 we have diam $x_n(\lambda) \leq 4\epsilon_\mu$, hence $| x_n(w_1') - x_n(w_2') | \leq$ diam $x_n(\lambda)$ $\leq 4\epsilon_\mu$. On the other hand $X_n(w_i') = x_{n\nu_i}(w_i') = x_n(w_i')$, $i = 1, 2$. Therefore

$$| X_n(w_1) - X_n(w_2) | \leq | X_n(w_1) - x_n(w_1') | + | x_n(w_1') - x_n(w_2') | + | x_n(w_2') - X_n(w_2) | \leq 2\epsilon_n + 4\epsilon_\mu + 2\epsilon_n \leq 8\epsilon_\mu < \epsilon.$$

c) $w_i \in \pi_i$, $\pi_i \in T_{n\nu_i}'$, $\pi_i \subset r_i$, $r_i \in \Delta_{n\mu}$, $i = 1, 2$. Here π_i is completely contained in r_i, hence $n \geq \nu_i \geq \mu$ (no. 37). Let w_{i,ν_i+1} be any point $w_{i,\nu_i+1} \in \pi_i^*$, hence $| X_n(w_i) - X_n(w_{i,\nu_i+1}) | < 2\epsilon_n$, $X_n(w_{i,\nu_i+1}) = x_n(w_{i,\nu_i+1})$, $i = 1, 2$. We have $\pi_i \in T_{n\nu_i}'$, hence $\pi_i \subset r_{i\nu_i}$, $r_{i\nu_i} \in \Delta_{n\nu_i}$. For any point $w_{i\nu_i} \in r_{i\nu_i}^*$ we have $| x_n(w_{i,\nu_i+1}) - x_n(w_{i\nu_i}) | \leq \operatorname{diam} \Sigma_{\nu_i}$, where Σ_{ν_i} is the surface defined by application of Lemma 19 on the rectangle $r_{i\nu_i}$, hence (no. 36) $| x_n(w_{i,\nu_i+1}) - x_n(w_{i,\nu_i}) | \leq \operatorname{diam} \Sigma_{\nu_i} < 30K\epsilon_{\nu_i}$. Now $\pi_i \in T_{n\nu_i}'$, hence no polygon $\pi \in T_{n1}' + \cdots + T_{n,\nu_i-1}'$ (no. 37) has points in common with π_i. Therefore $r_{i\nu_i}$ is not completely contained in any of the polygons $\pi \in T_{n1}' + \cdots + T_{n,\nu_i-1}'$. In particular, $r_{i\nu_i}$ is not completely contained in any of the polygons $\pi \in T_{n,\nu_i-1}'$. The polygons $\pi \in T_{n,\nu_i-1} - T_{n,\nu_i-1}'$ are either completely contained in rectangles $r \in \Delta_{nt}$, $t \geq \nu_i$, or completely contained in $T_{n1}' + \cdots + T_{n,\nu_i-1}'$. Therefore $r_{i\nu_i}^*$ cannot be completely contained in the interior of any such polygons π. In any case we can choose on $r_{i\nu_i}^*$ the point $w_{i\nu_i}$ not interior to any of the polygons $\pi \in T_{n,\nu_i-1}$. The point $w_{i\nu_i}$ now chosen belongs to a rectangle $r_{i,\nu_i-1} \in \Delta_{n,\nu_i-1}$, and for any point $w_{i,\nu_i-1} \in r_{i,\nu_i-1}^*$ we have $| x_n(w_{i,\nu_i}) - x_n(w_{i,\nu_i-1}) | \leq \operatorname{diam} \Sigma_{\nu_i-1} < 30K\epsilon_{\nu_j-1}$, where Σ_{ν_i-1} is defined as Σ_{ν_i}. By repeating this procedure we have a finite chain of points $w_{it} \in r_{it}^*$, $t = \nu_i, \nu_i - 1, \cdots, \mu + 1, \mu$, and $| x_n(w_{i,t+1}) - x_n(w_{it}) | < 30K\epsilon_t$, $t = \nu_i, \nu_i - 1, \cdots, \mu$, $i = 1, 2$. Therefore $| x_n(w_{i,\nu_i+1}) - x_n(w_{i\mu}) | \leq \Sigma | x_n(w_{it}) - x_n(w_{i,t+1}) | < 30K(\epsilon_\mu + \epsilon_{\mu+1} + \cdots + \epsilon_{\nu_i})$ and, because $\epsilon_{t+1} < \epsilon_t/2$, also $< 60K\epsilon_\mu$. Finally, if we denote by w_i', w_i'' the points w_{i,ν_i+1}, $w_{i\mu}$, we have

$$| X_n(w_i) - x_n(w_i') | < 2\epsilon_n \leqq 2\epsilon_\mu, \quad | x_n(w_i') - x_n(w_i'') | < 60K\epsilon_\mu, \quad w_i'' \in r_{i\mu}, \quad i = 1, 2.$$

Let us now observe that the points $w_{i,t+1}$, w_{it} belong to the same rectangle $r_{it} \in \Delta_{nt}$, $(t = \nu_i, \nu_i - 1, \cdots, \mu)$, hence the distance $| w_{i,t+1} - w_{it} |$ is $<$ twice the maximum dimension of all rectangles $r \in \Delta_{nt}$, i. e., by no. 30, $| w_{i,t+1} - w_{it} | < 4\xi_t$, where $4\xi_t \leq (1/2)\eta_{t-1}$, $(1/2^2)\eta_{t-2}, \cdots$. Therefore

$$| w_i' - w_{i,\mu+1} | = | w_{i,\nu_i+1} - w_{i,\mu+1} | \leq \Sigma | w_{i,t+1} - w_{it} |, \quad t = \nu_i, \nu_i - 1, \cdots, \mu + 1,$$

and finally

$$| w_i' - w_{i,\mu+1} | < 4\xi_{\nu_i} + 4\xi_{\nu_i-1} + \cdots + 4\xi_{\mu+1} \leq (1/2)\eta_\mu + (1/2^2)\eta_\mu + \cdots + (1/2^{\nu_i-\mu})\eta_\mu < \eta_\mu.$$

This assures that the point $w_{i,\mu+1}$ has a distance from $w_i' \in r_i[r_{i\mu}, r_i \in \Delta_{n\mu}]$ which is less than the minimum dimension of all rectangles $r \in \Delta_{n\mu}$. Also the point $w_i'' = w_{i\mu}$ belongs to the rectangle $r_i' \equiv r_{i\mu}$ and necessarily the rectangle r_i' is either coincident with r_i, or adjacent to r_i, or has a vertex in common with r_i, $i = 1, 2$. Therefore the rectangles r_1', r_2' belong to a chain r_1', r_1, r_2, r_2' of rectangles $r \in \Delta_{n\mu}$, any two consecutive ones being identical, or adjacent, or with a vertex in common. Consequently the points $w_1'' \in r_1'^*$, $w_2'' \in r_2'^*$ can be joined by a polygonal line λ formed by at most eight consecutive segments each contained in a side of a rectangle $r \in \Delta_{n\mu}$. This implies that

$$| x_n(w_1'') - x_n(w_2'') | \leq \operatorname{diam} x_n(\lambda) \leq 8\epsilon_\mu.$$

Finally we have

$$\begin{aligned} | X_n(w_1) - X_n(w_2) | &\leq | X_n(w_1) - x_n(w_1') | + | x_n(w_1') - x_n(w_1'') | \\ &+ | x_n(w_1'') - x_n(w_2'') | + | x_n(w_2'') - x_n(w_2') | + | x_n(w_2') - X_n(w_2) | \\ &\leq 2\epsilon_\mu + 60K\epsilon_\mu + 8\epsilon_\mu + 60K\epsilon_\mu + 2\epsilon_\mu \leq 132K\epsilon_\mu < \epsilon. \end{aligned}$$

d) $w_i \in Q - I_n'$, $i = 1, 2$. Here w_i not $\epsilon I_n'$, $I_n' = T_{n1}' + T_{n2}' + \cdots + T_{nn}'$, hence w_i not $\epsilon(T_{n1}' + \cdots + T_{n,n-1}')$. If w_i belonged to a polygon π, $\pi \in T_{nn} - T_{nn}'$ (as an interior or boundary point), then necessarily the polygon π would not be completely contained in the closed set $T_{n1}' + T_{n2}' + \cdots + T_{n,n-1}'$, hence $\pi \in T_{nn}'$ (no. 37), $\pi \subset I_n'$, which is impossible. This proves that w_i not ϵT_{nn}. Let us indicate by μ_i, $n \geq \mu_i \geq \mu$, the minimum of the indices t such that w_i not ϵT_{nt}, T_{nt} closed, $t = n, n-1, \cdots, \mu_i$. Then w_i not $\epsilon T_{n\mu_i}$, $w_i \in r_{i\mu_i}$, $i = 1, 2$. We have $X_n(w_i) = x_n(w_i)$, $i = 1, 2$. We can repeat the reasoning of c) with the simplification due to the fact that here $w_i' = w_i$, $i = 1, 2$. We obtain

$$\begin{aligned} | X_n(w_1) - X_n(w_2) | = | x_n(w_1) - x_n(w_2) | &\leq | x_n(w_1) - x_n(w_1'') | \\ &+ | x_n(w_1'') - x_n(w_2'') | + | x_n(w_2'') - x_n(w_2) | \\ &< 60K\epsilon_\mu + 4\epsilon_\mu + 60K\epsilon_\mu \leq 124K\epsilon_\mu < \epsilon. \end{aligned}$$

e) The two points w_i are in two different ones of the situations discussed in b), c), d). We have only to combine the different reasonings.

40. We have proved in the previous no. 39 that, given $\epsilon > 0$, there is an integer $\mu = \mu(\epsilon) > 0$ and a real number $\delta = \delta(\epsilon) > 0$ such that, for any two points $w_1, w_2 \in Q$, $| w_1 - w_2 | < \delta$, and for any integer $n \geq \mu$, we have

$|X_n(w_1) - X_n(w_2)| < \epsilon$. Now the $\mu - 1$ vectors $X_n(w)$, $w \in Q$, $n = 1, 2, \cdots, \mu - 1$, are continuous on Q, hence there is a number $\delta' = \delta'(\epsilon) > 0$ such that for any $w_1, w_2 \in Q$, $|w_1 - w_2| < \delta'$, and any $n = 1, 2, \cdots, \mu - 1$, we have $|X_n(w_1) - X_n(w_2)| < \epsilon$. If $\delta_0 = \delta_0(\epsilon) = \min[\delta, \delta']$, we have $|X_n(w_1) - X_n(w_2)| < \epsilon$ for any $w_1, w_2 \in Q$, $|w_1 - w_2| < \delta_0$ and any $n = 1, 2, \cdots,$. We have proved that the vectors $X_n(w)$, $w \in Q$, $n = 1, 2, \cdots$, are *equicontinuous* on Q. From $[\bar{S}_n] \subset A_{2\delta_n}$, $n = 1, 2, \cdots$ (no. 36), it follows that the vectors $X_n(w)$, $w \in Q$, $n = 1, 2, \cdots$, are also *equibounded* in Q.

41. By Ascoli's theorem there exists a subsequence $X_{n_m}(w)$ of the vectors $X_n(w)$, $n = 1, 2, \cdots$, which converges uniformly toward a continuous vector $X(w)$, $w \in Q$. For simplicity's sake we suppose $n_m = \overline{m}$. Such a vector coincides, on Q^*, with the vector $X(w)$ already determined in no. 29. Let S be the surface $S: x = X(w)$, $w \in Q$. From $X_n(w) \rightrightarrows X(w)$, $w \in Q$, and $\bar{S}_n: x = X_n(w)$, $w \in Q$, it follows that $\| \bar{S}_n, S \| \to 0$ and, by no. 29, also $\partial S = C$. We have $\| \bar{S}_n, S \| \to 0$ and, by no. 38, $a(\bar{S}_n) < (1/m)\mathfrak{J}(\bar{S}_n) < (1/m)(j+1) < +\infty$, hence $0 \leq L(S) \leq \lim a(\bar{S}_n) < +\infty$. From $[\bar{S}_n] \subset A_{2\delta_n}$, $\delta_n \to 0$, it follows that $[S] \subset A$, where A is a closed convex set of E_3. By no. 27 we have $S \subset W$, hence $\mathfrak{J}(S) \geq i$. Now we have $\delta_n \leq 1$, $n = 1, 2, \cdots$, hence all surfaces S, $\bar{S}_n$ are contained in the finite closed set A_2. We have $F_0(x, p) > 0$ for any $x \in A_2$, $|p| \neq 0$. The surface S is contained in A, hence in A_2, and the function $F_0(x, p)$ has continuous first derivatives with respect to p^s, $s = 1, 2, 3$, for all $x \in A$, $|p| \neq 0$ [because $F = F_0$ for $x \in A$], hence $F_0(x, p)$ has continuous first derivatives also for all $x \in [S]$, $|p| \neq 0$. We know that $\mathfrak{E}(x, p, \bar{p}) \geq 0$ for all $p \neq \bar{p}$, $|p|$, $|\bar{p}| \neq 0$, $x \in A$, hence also for all $x \in [S]$. All conditions of Lemma 9 (note in no. 8), where A is now the set A_2, are satisfied. Therefore the integral

$$\mathfrak{J}(S) = \int\int_S F_0(x, p)\, dw$$

is lower semicontinuous on the surface S with respect to the surfaces $\bar{S}_n$, $n = 1, 2, \cdots$, i. e. $\mathfrak{J}(S) \leq \underline{\lim}\, \mathfrak{J}(\bar{S}_n)$, where, by no. 38, we know that $\mathfrak{J}(\bar{S}_n) < j + 2^{-n}$, $n = 1, 2, \cdots$. Hence $\mathfrak{J}(S) \leq j$ and finally $i \leq \mathfrak{J}(S) \leq j$, where, by no. 27, we know that $j \leq i$. This implies that $\mathfrak{J}(S) = i = j$. We have thus proved that there exists a surface $S \in W$ for which $\mathfrak{J}(S)$ assumes its absolute minimum i in W. Theorem I (no. 25) is completely proved.

42. THEOREM II. *Let C be a closed Jordan curve in E_3, W the family of all oriented Fréchet surfaces S in E_3, whose Lebesgue area is finite and whose boundary curve is C, and let us suppose W is not empty. Let $\Im(S)$ be any positive definite semiregular integral such that $m\,|\,p\,| \leq F(x, p) \leq M\,|\,p\,|$, where $m, M, 0 < m \leq M < +\infty$, are two given constants, $x \in E_3$, $|\,p\,| \neq 0$. Then $\Im(S)$ has an absolute minimum in W.*

The proof is the same as for Theorem I with the obvious simplification in no. 26. Here we have $A = E_3$. From the equicontinuity of the vectors $X_n(w)$ in Q and the fact that we have $\| C_n, C \| \to 0$, it follows that the same vectors $X_n(w)$ are also equibounded in Q and therefore we can apply, here also, Ascoli's theorem.

43. *Note.* In the proofs of Theorems I and II the principle of Dirichlet is not applied directly. We observe however that Lemmas 5, 6, 12 have been applied, that the proof of Lemma 6 is based on Lemma 5 and that the recalled proofs of Lemmas 5 and 12 follow reasonings which are the same as for some existence theorems for the problem of Dirichlet. Nevertheless our proofs appear to need somewhat weaker facts than the principle of Dirichlet, namely only that a) any non-degenerate surface of finite Lebesgue area posseses Dirichlet representations; b) the g. l. b. of the relative Dirichlet integrals coincides with the Lebesgue area, without stating that the minimum is actually attained.

THE UNIVERSITY OF BOLOGNA AND
PURDUE UNIVERSITY.

BIBLIOGRAPHY.

[1] L. Cesari, "Criteri di uguale continuità ed applicazioni alla quadratura delle superficie," *Annali Scuola Normale Superiore Pisa*, vol. 12 (1943), pp. 61-84.

[2] ———, "La nozione di integrale sopra una superficie in forma parametrica," *Ibid.*, vol. 13 (1947), pp. 73-117.

[3] ———, "Condizioni sufficienti per la semicontinuità degli integrali sopra una superficie in forma parametrica," *Ibid.*, vol. 14 (1948), pp. 47-79.

[4] ———, "Condizioni necessarie per la semicontinuità degli integrali sopra una superficie in forma parametrica," *Annali di Matematica*, vol. 29 (1950), pp. 199-224.

[5] ———, "Area and representation of surfaces," *Bulletin of the American Mathematical Society*, vol. 56 (1950), pp. 218-232.

[6] ———, "Sulle superficie di Fréchet," *Rivista di Matematica dell'Università di Parma*, vol. 1 (1950), pp. 19-44.

[7] ———, "Su un particolare processo di retrazione per superficie," *Ibid.*, vol. 2 (1951).

[8] ———, "Problemi di Calcolo delle Variazioni e questioni connesse," *Ibid.*, vol. 1 (1950), pp. 293-304.

[9] S. Cinquini, "Sull'approssimazione delle funzioni di due variabili," *Annali di Matematica*, vol. 11 (1933), pp. 295-323.

[10] R. Courant, "On the problem of Plateau," *Proceedings of the National Academy of Science*, vol. 22 (1936), pp. 367-372.

[11] J. Douglas, "Solution of the problem of Plateau," *Transactions of the American Mathematical Society*, vol. 33 (1931), pp. 263-321.

[12] S. Eilenberg, "On Φ measures," *Annales Société Polonaise de Mathématique*, vol. 17 (1938), pp. 252-253.

[13] ———, (with O. E. Harrold) "Continua of finite linear measure," *American Journal of Mathematics*, vol. 65 (1943), pp. 137-146.

[14] E. J. McShane, "On the semicontinuity of integrals in the Calculus of Variations," *Annals of Mathematics*, vol. 33 (1932), pp. 460-484.

[15] ———, "Integrals over surfaces in parametric form," *Annals of Mathematics*, vol. 34 (1933), pp. 815-838.

[16] ———, "Parametrizations of saddle surfaces, with applications to the problem of Plateau," *Transactions of the American Mathematical Society*, vol. 35 (1933), pp. 716-733.

[17] ———, "Existence theorems for double integral problems of the Calculus of Variations," *Ibid.*, vol. 38 (1935), pp. 549-563.

[18] ———, "Extension of range of functions," *Bulletin of the American Mathematical Society*, vol. 40 (1934), pp. 837-842.

[19] C. B. Morrey, "A class of representation of manifolds," *American Journal of Mathematics*, vol. 55 (1933), pp. 683-707; vol. 56 (1934), pp. 275-293.

[20] ———, "The topology of (path) surfaces," *Ibid.*, vol. 57 (1935), pp. 17-50.

[21] ———, "An analytic characterization of surfaces of finite Lebesgue area," *Ibid.*, vol. 57 (1935), pp. 692-702; vol. 58 (1936), pp. 313-322.

[22] T. Rado, "On the problem of Plateau," *Ergebnisse der Mathematik und ihrer Grenzgebiete*, vol. 2 (1933).

[23] ———, "On the semicontinuity of double integrals in the parametric form," *Transactions of the American Mathematical Society*, vol. 51 (1942), pp. 336-361.

[24] ———, "Length and Area," *Colloquium Publications of the American Mathematical Society*, vol. 30 (1948).

[25] P. V. Reichelderfer, "On the semicontinuity of double integrals," *Transactions of the American Mathematical Society*, vol. 67 (1949), pp. 275-285.

[26] L. Tonelli, *Fondamenti di Calcolo delle Variazioni*, Bologna, Zanichelli (1923-1924).

[27] ———, "Sulla rappresentazione analitica delle funzioni di variabili reali," *Rendiconti Circolo Matematico di Palermo*, vol. 29 (1910), pp. 1-36.

[28] ———, "Sur la semicontinuité des integrales doubles du Calcul des Variations," *Acta Mathematica*, vol. 53 (1929), pp. 325-346.

[29] ———, "L'estremo assoluto degli integrali doppi," *Annali Scuola Normale Superiore Pisa*, vol. 2 (1933), pp. 89-130.

[30] ———, "Su l'esistenze delle estremali assolute per gli integrali doppi," *Ibid.*, vol. 8 (1939), pp. 161-165.

[31] ———, "Sul problema di Plateau," *Rendiconti Accademia dei Lincei*, vol. 29 (1936), pp. 333-339, 393-398.

[32] L. C. Young, "A lemma in the theory of surfaces," *Journal of the London Mathematical Society*, vol. 19 (1944), pp. 209-219.

[33] ———, "On the isoperimetric ratio for a harmonic surface," *Proceedings of the London Mathematical Society*, vol. 49 (1947), pp. 396-408.

[34] ———, "On area and length," *Fundamenta Mathematicae*, vol. 35 (1948), pp. 275-302.

[35] ———, "Some applications of the Dirichlet integral to the theory of surfaces," *Transactions of the American Mathematical Society*, vol. 64 (1948), pp. 317-335.

THE "INDEPENDENT SCALARS" IN HOMOGENEOUS TURBULENCE.*

By N. Coburn.

1. Introduction. In discussing my paper[1] at the University of Illinois Symposium on Fluid Dynamics, Professor S. Chandrasekhar mentioned the problem of determining a method for generating the "independent scalars" in various types of homogeneous turbulence. By definition, these scalars must possess three properties: (1) they must be components of the various correlation tensors; (2) these scalars must be defined in such a manner that the correlation tensors satisfy the continuity relations identically; (3) the maximum number of such scalars must occur in each correlation tensor.

It is the purpose of this paper to furnish a method for the construction of such independent scalars. The method is based on the following considerations. First, one determines the maximum number of independent scalars associated with a given correlation tensor. Secondly, by a contraction process, a new tensor, containing the desired number of independent scalars, is constructed. By forming the curl of this tensor, one obtains a tensor which has two properties: (1) its divergence is zero (and hence it satisfies the continuity relations); (2) it possesses the structure of a correlation tensor which depends upon the maximum number of independent scalars.

In order to develop the method, we consider the cases of: (1) isotropic turbulence; (2) turbulence possessing axial symmetry with respect to a specified direction. The expressions for the correlation tensors in terms of the independent scalars are well known for the first case from the work of Kármán-Howarth [2] and H. P. Robertson.[3] For the second case, S. Chandrasekhar [4] has derived, recently, the corresponding results for the correlation tensors.

---

* Received October 26, 1950.

[1] N. Coburn, "A method for constructing correlation tensors in homogeneous turbulence," *Proceedings of the University of Illinois Symposium on Fluid Dynamics*, to be published.

[2] T. von Kármán and L. Howarth, "Isotropic turbulence," *Proceedings of the Royal Society*, vol. 164 (1938), pp. 192-215.

[3] H. P. Robertson, "An invariant theory of isotropic turbulence," *Proceedings of the Cambridge Philosophical Society*, vol. 36 (1940), pp. 209-223.

[4] S. Chandrasekhar, "The theory of axissymmetric turbulence," *Philosophical Transactions of the Royal Society*, vol. 242 (1950), pp. 557-577.

Finally, we discuss the general theory for constructing correlation tensors in terms of independent scalars. It is shown that if we require the correlation tensors to be *linear functions* of the independent scalars and their derivatives then, essentially, the above method is the one and only method for constructing such tensors.

The author wishes to thank Professor Chandrasekhar for allowing the author to see the proofs of his papers on axisymmetric turbulence (see footnote 4) and for his remarks relating the results of this paper with those of his paper.

2. The independent scalars in isotropic turbulence. Here, we consider the theory of Kármán-Howarth and Robertson. From the work of either of these authors, it is known that the correlation tensors have simple forms. In order to introduce these formulas, we mention our notation. Let x^λ, $\lambda = 1, 2, 3$, denote a Cartesian orthogonal coordinate system in Euclidean three-space. Further, if $P(x^\lambda)$, $'P('x^\lambda)$ denote the two points involved in the construction of the correlation tensors, then $\xi^\lambda = {}'x^\lambda - x^\lambda$ will denote the vector joining these points. Evidently, the magnitude, r, of this vector is,

$$r = [(\xi^1)^2 + (\xi^2)^2 + (\xi^3)^2]^{1/2}.$$

One additional tensor, which enters into this theory, is the metric tensor, $g^{\lambda\mu}$. With the aid of this tensor and vector, ξ^λ, the correlation tensors may be written as

$$R^\lambda = A\xi^\lambda, \qquad R^{\lambda\mu} = B\xi^\lambda\xi^\mu + Cg^{\lambda\mu},$$

$$R^{\lambda\mu\alpha} = D\xi^\lambda\xi^\mu\xi^\alpha + E(\xi^\lambda g^{\mu\alpha} + \xi^\mu g^{\lambda\alpha}) + F\xi^\alpha g^{\lambda\mu}, \tag{2.1}$$

where A, B, C, D, E, F are scalar functions of r.

These tensors satisfy the following continuity relations

$$R^\lambda{}_{,\lambda} = R^{\lambda\mu}{}_{,\mu} = R^{\lambda\mu\alpha}{}_{,\alpha} = 0, \tag{2.2}$$

where commas denote differentiation, that is $R^{\lambda\mu}{}_{,\mu} = \partial R^{\lambda\mu}/\partial\xi^\mu$ etc. The relation (2.2) furnishes a first order differential equation for A. Hence R^λ does not depend upon an independent scalar. Further, we notice that (2.2) determines a vanishing correlation vector and a vanishing correlation tensor of the second order. That is, the above relations imply that B, C are related by a single first order differential equation and that D, E, F are related by two first order differential equations. Thus, both $R^{\lambda\mu}$ and $R^{\lambda\mu\alpha}$ depend upon one independent scalar. In the next two paragraphs, we shall determine expressions for $R^{\lambda\mu}$, $R^{\lambda\mu\alpha}$ which depend upon single independent scalars, respectively.

First, we consider the tensor $R^{\lambda\mu}$. Multiplying the second equation of (2.1) by $e_{\gamma\lambda\beta}\xi^{\beta}$, we obtain

$$\bar{S}_{\gamma}{}^{\cdot\mu} = e_{\gamma\lambda\beta}R^{\mu\lambda}\xi^{\beta} = Ce_{\gamma\lambda\beta}g^{\lambda\mu}\xi^{\beta}. \tag{2.3}$$

Note, $e_{\gamma\lambda\beta}$ is the permutation tensor density. If we restrict our coordinate systems to right handed Cartesian orthogonal coordinates then $e_{\gamma\lambda\beta}$ is a tensor. Next, we introduce the tensor

$$\bar{R}^{\mu\alpha} = e^{\alpha\gamma\rho}\bar{S}_{\gamma}{}^{\cdot\mu}{}_{,\rho}. \tag{2.4}$$

It is easily verified that the tensor $\bar{R}^{\mu\alpha}$ satisfies (2.2), identically. That is, this tensor has zero divergence. Further, from (2.3), we see that this tensor depends upon one scalar. It remains to be shown that $\bar{R}^{\mu\alpha}$ has the form (2.1) and hence can be considered as a correlation tensor. Substituting (2.3) into (2.4), we obtain

$$\bar{R}^{\mu\alpha} = -\delta_{\lambda\beta}{}^{\alpha\rho}g^{\lambda\mu}(C\delta_{\rho}{}^{\beta} + r^{-1}\xi^{\beta}\xi^{\gamma}g_{\rho\gamma}dC/dr), \tag{2.5}$$

where $\delta_{\beta}{}^{\alpha}$ is the Kronecker tensor and $\delta_{\lambda\beta}{}^{\alpha\rho}$ is the generalized Kronecker tensor. In particular, $\delta_{\lambda\beta}{}^{\alpha\rho}$ satisfies the equation

$$\delta_{\lambda\beta}{}^{\alpha\rho} = e^{\mu\alpha\rho}e_{\mu\lambda\beta} = \delta_{\lambda}{}^{\alpha}\delta_{\beta}{}^{\rho} - \delta_{\lambda}{}^{\rho}\delta_{\beta}{}^{\alpha}.$$

Through use of this relation, we may reduce (2.5) to the equation

$$\bar{R}^{\mu\alpha} = -(2C + rdC/dr)g^{\mu\alpha} + r^{-1}dC/dr\xi^{\alpha}\xi^{\mu}.$$

That is, the above expression furnishes *a second order correlation tensor which depends upon a single scalar and for which the continuity relation* (2.2) *is identically satisfied.* It should be noted that (2.5) does not assure us that $\bar{R}^{\mu\alpha}$ will be symmetric. Such symmetry can be obtained, if needed, by use of an additional procedure, to be explained in section 6.

Now, we turn to the study of the third order correlation tensor, $R^{\lambda\mu\alpha}$. Similarly to (2.3) we introduce the tensor

$$\bar{S}_{\gamma}{}^{\cdot\lambda\mu} = e_{\gamma\alpha\beta}R^{\lambda\mu\alpha}\xi^{\beta} = Ee_{\gamma\alpha\beta}(\xi^{\lambda}g^{\mu\alpha} + \xi^{\mu}g^{\lambda\alpha})\xi^{\beta}. \tag{2.6}$$

Further, we construct the tensor

$$\bar{R}^{\lambda\mu\nu} = e^{\nu\gamma\rho}\bar{S}_{\gamma}{}^{\cdot\lambda\mu}{}_{,\rho}. \tag{2.7}$$

Evidently, the divergence relation (2.2) is identically satisfied. Further, the tensor $\bar{R}^{\lambda\mu\alpha}$ depends upon a single scalar. Hence this scalar is an independent scalar. Substituting (2.6) into (2.7), we obtain

(2.8)
$$\bar{R}^{\lambda\mu\nu} = -\delta_{\alpha\beta}{}^{\nu\rho}[r^{-1}dE/dr g_{\delta\rho}\xi^{\delta}\xi^{\beta}(\xi^{\lambda}g^{\mu\alpha} + \xi^{\mu}g^{\lambda\alpha}) + E(\delta_{\rho}{}^{\lambda}g^{\mu\alpha}\xi^{\beta} + \delta_{\rho}{}^{\mu}g^{\lambda\alpha}\xi^{\beta} + \delta_{\rho}{}^{\beta}\xi^{\lambda}g^{\mu\alpha} + \delta_{\rho}{}^{\beta}\xi^{\mu}g^{\lambda\alpha})].$$

Simplifying (2.8), we find that

(2.9) $$\bar{R}^{\lambda\mu\nu} = 2(r^{-1}dE/dr)\xi^{\lambda}\xi^{\mu}\xi^{\nu} - (3E + rdE/dr)(\xi^{\lambda}g^{\mu\nu} + \xi^{\mu}g^{\lambda\nu}) + 2E\xi^{\nu}g^{\lambda\mu}.$$

Thus, (2.9) *furnishes a third order correlation tensor which depends upon a single scalar and for which the continuity relation* (2.2) *is identically satisfied. That is, we have constructed a third order correlation tensor of the type given in* (2.1), *and such that this tensor is expressed in terms of one independent scalar* (*the maximum number of such scalars*).

3. The independent scalar in turbulence with axial symmetry. The general theory of turbulence with axial symmetry has been initiatel by G. K. Batchelor [5] and completely studied by S. Chandrasekhar.[4] In addition, another approach to this theory has been given by the present author.[1] The previously mentioned authors have followed the invariant theory method of H. P. Robertson.[3] Here, we shall apply our present ideas to the work of S. Chandrasekhar.

Let j^{λ} $(\lambda = 1, 2, 3)$ denote a unit vector in the axial direction. By following the invariant theory approach, Chandrasekhar constructs the following correlation tensors:

(3.1) $$R^{\lambda} = M\xi^{\lambda} + Nj^{\lambda},$$

(3.2) $$R^{\lambda\mu} = A\xi^{\lambda}\xi^{\mu} + Bg^{\lambda\mu} + Cj^{\lambda}j^{\mu} + Dj^{\lambda}\xi^{\mu} + E\xi^{\lambda}j^{\mu},$$

(3.3) $$R^{\lambda\mu\alpha} = F\xi^{\lambda}\xi^{\mu}\xi^{\alpha} + G(j^{\lambda}\xi^{\mu} + \xi^{\lambda}j^{\mu})\xi^{\alpha} + H(j^{\lambda}\xi^{\mu} + \xi^{\lambda}j^{\mu})j^{\alpha} + Ig^{\lambda\mu}\xi^{\alpha} + Jg^{\lambda\mu}j^{\alpha} + Kj^{\lambda}j^{\mu}j^{\alpha} + Lj^{\lambda}j^{\mu}\xi^{\alpha} + Q\xi^{\mu}\xi^{\lambda}j^{\alpha} + R(\xi^{\lambda}g^{\mu\alpha} + \xi^{\mu}g^{\lambda\alpha}) + S(j^{\lambda}g^{\mu\alpha} + j^{\mu}g^{\lambda\alpha}).$$

The scalars N, M, etc., are considered to be functions of r, and μ where

(3.4) $$r = [(\xi^1)^2 + (\xi^2)^2 + (\xi^3)^2]^{1/2}, \qquad r\mu = \xi^{\lambda}j_{\lambda}.$$

The problem is to find the correlation tensors in terms of the independent scalars for which the continuity relations (2.2) will be identically satisfied. First, we note that (2.2) furnishes a single scalar relation between the scalars

[5] G. K. Batchelor, "On the theory of axisymmetric turbulence," *Proceedings of the Royal Society*, vol. 186 (1946), pp. 480-490.

M, N of (3.1). Hence, only one of the scalars, M, N, is independent. Secondly, we note that the second continuity relation determines a correlation vector. Since the general correlation vector for axial symmetry is of the type (3.1), the equation (2.2) furnishes two relations for the coefficients A, B, C, D, E. That is, three of these scalars are independent. By a similar argument, we see that the ten coefficients F, G, etc., satisfy four relations. Hence, six independent scalars characterize the third order correlation tensor $R^{\lambda\mu\alpha}$.

In order to express the correlation vector, R^{λ}, in terms of a single independent scalar, we proceed as in section 2. Multiplying (3.1) by $e_{\gamma\lambda\beta}\xi^{\beta}$ we obtain

$$\bar{S}_{\gamma} = e_{\gamma\lambda\beta}R^{\lambda}\xi^{\beta} = Ne_{\gamma\lambda\beta}j^{\lambda}\xi^{\beta}. \tag{3.5}$$

We shall show that the vector

$$\bar{R}^{\alpha} = e^{\alpha\gamma\rho}\bar{S}_{\gamma,\rho} \tag{3.6}$$

is the desired vector. First, we note that the vector, $\bar{R}^{\alpha}$, has zero divergence. Secondly, we note that

$$\partial/\partial\xi^{\rho} = g_{\rho\nu}[\xi^{\nu}(r^{-1}\partial/\partial r - \mu r^{-2}\partial/\partial\mu) + j^{\nu}r^{-1}\partial/\partial\mu]. \tag{3.7}$$

Substituting (3.5) into (3.6), we obtain the formula

$$\bar{R}^{\alpha} = -\delta_{\lambda\beta}{}^{\alpha\rho}(j^{\lambda}\xi^{\beta}N_{,\rho} + Nj^{\lambda}\delta_{\rho}{}^{\beta}). \tag{3.8}$$

Simplifying (3.8) by use of the expansion formula for $\delta_{\lambda\mu}{}^{\alpha\rho}$ and (3.7) we find that

$$\bar{R}^{\alpha} = [\mu\partial N/\partial r + r^{-2}(1-\mu^{2})\partial N/\partial\mu]\xi^{\alpha} - [r\partial N/\partial r + 2N]j^{\alpha}. \tag{3.9}$$

The above equation is essentially equation (25) in Chandrasekhar's paper.[4]

Next, we consider the tensor $R^{\lambda\mu}$. We noted in the previous paragraph that this tensor should be expressible in terms of three independent scalars. Hence, we repeat the technique used in section 2. That is, we introduce the tensor

$$\bar{S}_{\gamma}{}^{\cdot\mu} = e_{\gamma\lambda\beta}R^{\lambda\mu}\xi^{\beta} = e_{\gamma\lambda\beta}(Bg^{\lambda\mu} + Cj^{\lambda}j^{\mu} + Dj^{\lambda}\xi^{\mu})\xi^{\beta}. \tag{3.10}$$

With the aid of (3.10), we form the tensor $\bar{R}^{\mu\alpha}$ of (2.4) and obtain

$$\begin{aligned}\bar{R}^{\mu\alpha} = &-\delta_{\lambda\beta}{}^{\alpha\rho}\xi^{\beta}[g^{\lambda\mu}(r^{-1}\partial B/\partial r\xi_{\rho} - r^{-2}\mu\partial B/\partial\mu\xi_{\rho} + r^{-1}\partial B/\partial\mu j_{\rho})\\ &+ j^{\lambda}j^{\mu}(r^{-1}\partial C/\partial r\xi_{\rho} - r^{-2}\mu\partial C/\partial\mu\xi_{\rho} + r^{-1}\partial C/\partial\mu j_{\rho})\\ &+ j^{\lambda}\xi^{\mu}(r^{-1}\partial D/\partial r\xi_{\rho} - r^{-2}\mu\partial D/\partial\mu\xi_{\rho} + r^{-1}\partial D/\partial\mu j_{\rho}) + Dj^{\lambda}\delta_{\rho}{}^{\mu}]\\ &-\delta_{\lambda\beta}{}^{\alpha\rho}\delta_{\rho}{}^{\beta}[Bg^{\lambda\mu} + Cj^{\lambda}j^{\mu} + Dj^{\lambda}\xi^{\mu}].\end{aligned} \tag{3.11}$$

For the sake of simplicity, we have written ξ_ρ for $g_{\rho\gamma}\xi^\gamma$ in the above equation. Replacing $\delta_{\lambda\beta}{}^{\alpha\rho}$ by $(\delta_\lambda{}^\alpha\delta_\beta{}^\rho - \delta_\beta{}^\alpha\delta_\lambda{}^\rho)$ we find that (3.11) becomes

(3.12)
$$\begin{aligned}\bar{R}^{\mu\alpha} = {} & \xi^\alpha\xi^\mu[r^{-1}\partial B/\partial r - r^{-2}\mu\partial B/\partial\mu + \mu\partial D/dr + (1-\mu^2)\partial D/\partial\mu] \\ & - g^{\alpha\mu}(2B + r\partial B/\partial r) + \xi^\alpha j^\mu[r^{-1}\partial B/\partial\mu + \mu\partial C/\partial r \\ & + r^{-1}(1-\mu^2)\partial C/\partial\mu + D] - j^\alpha\xi^\mu(r\partial D/\partial r + 3D) \\ & - j^\alpha j^\mu(2C + r\partial C/\partial r).\end{aligned}$$

The tensor $\bar{R}^{\mu\alpha}$ of (3.12) does not coincide with the corresponding tensor developed by Chandrasekhar.[6] The reason for this difference lies in the non-uniqueness of A, B, etc. That is, another possible correlation tensor can be formed from linear combinations of A, B, etc. or from

$$'\bar{S}_\gamma{}^{\cdot\mu} = e_{\gamma\lambda\beta}{}'R^{\lambda\mu}j^\beta = e_{\gamma\lambda\beta}(F\xi^\lambda\xi^\mu + Gg^{\lambda\mu} + H\xi^\lambda j^\mu)j^\beta, \tag{3.13}$$

where $'R^{\lambda\mu} = F\xi^\lambda\xi^\mu + Gg^{\lambda\mu} + H\xi^\lambda j^\mu + Lj^\lambda\xi^\mu + Mj^\lambda j^\mu$. By expressing F, G, H as linear combinations of B, C, D with variable coefficients and adding the resulting tensor $'\bar{R}^{\lambda\mu}$ to $\bar{R}^{\lambda\mu}$, one obtains other possible second order correlation tensors. The uniqueness problem will be discussed in the next section. To calculate the third order correlation tensors, one applies (2.6), (2.7) to the tensor (3.3). The calculations are routine but rather lengthy and will be omitted.

4. The theory of correlation tensors which depend upon independent scalars. In this section, we discuss the mathematical significance of our method.

a. *The General Problem.* We consider the following problem: *given the arbitrary correlation tensors* $R^\lambda(\xi^\gamma)$, $R^{\lambda\mu}(\xi^\gamma)$, $R^{\lambda\mu\alpha}(\xi^\gamma)$, *and the independent fields* ξ^λ, j^λ, i^λ, *respectively; to determine associated tensors* $\bar{R}^\lambda(\xi^\gamma)$, $\bar{R}^{\lambda\mu}(\xi^\gamma)$, $\bar{R}^{\lambda\mu\alpha}(\xi^\gamma)$ *such that*:

(1) $\bar{R}^\lambda$ is a linear function of R^λ and its first derivatives with respect

---

[6] Professor Chanrasekhar has called my attention to the fact that the relations between his scalars Q_1, Q_2, Q_3 and the scalars B, C, D of (3.12) and F, G, H of (3.13) are

$$\begin{aligned}Q_1 &= B + C + r\mu D = -r\mu F - H, \\ Q_2 &= -C - r^{-1}\partial(r\mu C + r^2 D)/\partial\mu = H + r^{-1}\partial(r^3 F + r\mu H + G)/\partial\mu, \\ Q_3 &= -D - (r^{-1}\partial/\partial r - \mu r^{-2}\partial/\partial\mu)(r\mu C + r^2 D) \\ &= F + (r^{-1}\partial/\partial r - \mu r^{-2}\partial/\partial\mu)(r^2 F + G + r\mu H).\end{aligned}$$

to ξ^λ and is homogeneous in ξ^λ, j^λ and i^λ of order zero (and similarly for $\bar{R}^{\lambda\mu}$, $\bar{R}^{\lambda\mu\alpha}$ in terms of $R^{\lambda\mu}$, $R^{\lambda\mu\alpha}$, respectively);

(2) $\bar{R}^\lambda$ has identically vanishing divergence (and similarly for $\bar{R}^{\lambda\mu}$, $\bar{R}^{\lambda\mu\alpha}$);

(3) $\bar{R}^\lambda$, $\bar{R}^{\lambda\mu}$, $\bar{R}^{\lambda\mu\alpha}$ have the structure of correlation tensors.

We shall show that *the tensors* $\bar{R}^\lambda$, $\bar{R}^{\lambda\mu}$, $\bar{R}^{\lambda\mu\alpha}$ *formed in accord with the following rules and only in accord with these rules satisfy the conditions* (1) *through* (3).

(4) Construct the tensors $\bar{S}_\gamma$, $\bar{S}_\gamma{}^{\cdot\mu}$, $\bar{S}_\gamma{}^{\cdot\lambda\mu}$ from the given correlation tensors R^λ, $R^{\lambda\mu}$, $R^{\lambda\mu\alpha}$ and the independent vector fields ξ^λ, j^λ, i^λ according to the laws

$$(4.1)\qquad \bar{S}_\gamma = e_{\gamma\lambda\beta}R^\lambda\xi^\beta, \qquad {}'\bar{S}_\gamma = e_{\gamma\lambda\beta}{}'R^\lambda j^\beta, \qquad {}''\bar{S}_\gamma = e_{\gamma\lambda\beta}{}''R^\lambda i^\beta,$$

$$(4.2)\qquad \bar{S}_\gamma{}^{\cdot\mu} = e_{\gamma\lambda\beta}R^{\mu\lambda}\xi^\beta, \qquad {}'\bar{S}_\gamma{}^{\cdot\mu} = e_{\gamma\lambda\beta}{}'R^{\mu\lambda}j^\beta, \qquad {}''\bar{S}_\gamma{}^{\cdot\mu} = e_{\gamma\lambda\beta}{}''R^{\mu\lambda}i^\beta,$$

$$(4.3)\qquad \bar{S}_\gamma{}^{\cdot\lambda\mu} = e_{\gamma\alpha\beta}R^{\lambda\mu\alpha}\xi^\beta, \qquad {}'\bar{S}_\gamma{}^{\cdot\lambda\mu} = e_{\lambda\alpha\beta}{}'R^{\lambda\mu\alpha}j^\beta, \qquad {}''\bar{S}_\gamma{}^{\cdot\lambda\mu} = e_{\gamma\alpha\beta}{}''R^{\lambda\mu\alpha}i^\beta.$$

(5) The desired tensors, $\bar{R}^\lambda$, $\bar{R}^{\lambda\mu}$, $\bar{R}^{\lambda\mu\alpha}$, are determined by

$$(4.4)\qquad \bar{R}^\alpha = e^{\alpha\gamma\rho}\bar{S}_{\gamma,\rho}, \qquad {}'\bar{R}^\alpha = e^{\alpha\gamma\rho}{}'\bar{S}_{\gamma,\rho}, \qquad {}''\bar{R}^\alpha = e^{\alpha\gamma\rho}{}''\bar{S}_{\gamma,\rho},$$

$$(4.5)\qquad \bar{R}^{\mu\alpha} = e^{\alpha\gamma\rho}\bar{S}_\gamma{}^{\cdot\mu}{}_{,\rho}, \qquad {}'\bar{R}^{\mu\alpha} = e^{\alpha\gamma\rho}{}'\bar{S}_\gamma{}^{\cdot\mu}{}_{,\rho}, \qquad {}''\bar{R}^{\mu\alpha} = e^{\alpha\gamma\rho}{}''\bar{S}_\gamma{}^{\cdot\mu}{}_{,\rho},$$

$$(4.6)\qquad \bar{R}^{\lambda\mu\alpha} = e^{\alpha\gamma\rho}\bar{S}_\gamma{}^{\cdot\lambda\mu}{}_{,\rho}, \qquad {}'\bar{R}^{\lambda\mu\alpha} = e^{\alpha\gamma\rho}{}'\bar{S}_\gamma{}^{\cdot\lambda\mu}{}_{,\rho}, \qquad {}''\bar{R}^{\lambda\mu\alpha} = e^{\alpha\gamma\rho}{}''\bar{S}_\gamma{}^{\cdot\lambda\mu}{}_{,\rho}.$$

(6) The most general tensors of the desired type are of the form

$$(4.7)\qquad \bar{\bar{R}}^\lambda = a\bar{R}^\lambda + b'\bar{R}^\lambda + c''\bar{R}^\lambda,$$

$$(4.8)\qquad \bar{\bar{R}}^{\lambda\mu} = d\bar{R}^{\lambda\mu} + e'\bar{R}^{\lambda\mu} + f''\bar{R}^{\lambda\mu},$$

$$(4.9)\qquad \bar{\bar{R}}^{\lambda\mu\alpha} = g\bar{R}^{\lambda\mu\alpha} + h'\bar{R}^{\lambda\mu\alpha} + k''\bar{R}^{\lambda\mu\alpha},$$

where $a, b, c, \cdots$ are constants. The scalars of ${}'\bar{R}^\lambda$, ${}''\bar{R}^\lambda$ will be independent scalars if the scalars of $\bar{R}^\lambda$ are less than the maximum number of independent scalars. Otherwise, the scalars of ${}'\bar{R}^\lambda$, ${}''\bar{R}^\lambda$ will be linear combinations, with variable coefficients, of the scalars of $\bar{R}^\lambda$ (and similarly for the second and third order tensors).

b. *Remarks on the Above Problem.* We consider each of the statements (1) through (6) of section (a) and discuss these statements under the corresponding numbers of this section.

(1) This first statement implies that the scalars of $\bar{R}^\lambda$ are linear func-

tions of the scalars of R^λ and of the derivatives of these scalars. This will be shown in our proof (similar results apply to $\bar{R}^{\lambda\mu}$, $\bar{R}^{\lambda\mu a}$).

(2) The second statement implies that $\bar{R}^\lambda$, $\bar{R}^{\lambda\mu}$, $\bar{R}^{\lambda\mu a}$ are each expressed in terms of independent scalars. In this connection, anothed problem presents itself. Do $\bar{\bar{R}}^\lambda$, $\bar{\bar{R}}^{\lambda\mu}$, $\bar{\bar{R}}^{\lambda\mu a}$ contain the *maximum* number of scalars? We shall discuss this in (6).

(3) It is, of course essential that $\bar{R}^\lambda$, $\bar{R}^{\lambda\mu}$, $\bar{R}^{\lambda\mu a}$ etc., shall have the structure of correlation tensors. This result will follow from the fact that R^λ, $R^{\lambda\mu}$, $R^{\lambda\mu a}$ will be chosen to have such structure and the assumption (1). Again, this point must be verified in our proof.

(4) There can be at most three independent vector fields associated with three-space, independent of the type of turbulence. Of course, if the turbulence involves only one vector field, the primed and double primed tensors do not appear (see section 2). In axial symmetric turbulence, the double primed tensors do not occur, but the unprimed and primed tensors do occur. The tensors R^λ, $'R^\lambda$, $''R^\lambda$ (and similarly for the second and third order tensors) are assumed to be independent. Note, the equations (4. 1) through (4. 3) annul the non-independent scalars.

(5) It should be noted that the divergence of the tensors $\bar{R}^\lambda$, $\bar{R}^{\lambda\mu}$, $\bar{R}^{\lambda\mu a}$, etc., vanish identically in virtue of (4. 4) through (4. 6).

(6) We shall show that the process of section (a) generates *all correlation tensors* which depend *linearly* upon independent scalars and their derivatives and whose *divergence vanishes* when the tensors R^λ, $'R^\lambda$, $''R^\lambda$ (and those of the second and third order), are properly chosen. Then, the $\bar{\bar{R}}^\lambda$, $\bar{\bar{R}}^{\lambda\mu}$, $\bar{\bar{R}}^{\lambda\mu a}$ will contain the maximum number of independent scalars. The tensors R^λ, $'R^\lambda$, $''R^\lambda$ (and similarly for the corresponding second and third order tensors) are to be chosen as follows. Construct the tensor R^λ so that this tensor contains the maximum number of *distinct* scalars, N, not all of which are actually independent. Determine the number of conditions imposed on these scalars by the continuity relations. By subtracting the number of conditions from the maximum number of distinct scalars, one obtains *the maximum number of independent scalars, say m.* If $\bar{R}^\lambda$ contains $m_1 < m$ independent scalars, then $'\bar{R}^\lambda$ must be constructed. To do this, we introduce $'R^\lambda$. This tensor has the same structure as R^λ but depends upon N new distinct scalars. Thus, we construct a tensor $'\bar{R}^\lambda$ which depends upon m_2 (where $m_2 \leqq m - m_1$)

independent scalars. Finally, we construct $''\bar{R}^\lambda$ which depends upon m_3 (where $m_3 = m - m_1 - m_2$) independent scalars. Note, the tensor $'\bar{R}^\lambda$ may *appear* to contain more than $m - m_1$ independent scalars (see 3.12). In this case, the remaining scalars are linear combinations of the m independent scalars.

5. Proof of the results.

a. *The Tensors with Identically Vanishing Divergence which Depend Linearly upon R^λ, $R^{\lambda\mu}$, $R^{\lambda\mu\alpha}$ and Their Derivatives.* We consider the situation where two independent vector fields ξ^λ, j^λ exist. The case, in which i^λ exists, is treated in the same manner. First, consider the vector $\bar{R}^\lambda$. This vector is a general linear combination of R^λ and its derivatives with respect to ξ^λ. Hence, we may write

$$\begin{aligned}(5.1)\qquad \bar{R}^\lambda = {} & AR^\lambda + B\xi^\alpha R^\lambda{}_{,\alpha} + Cj^\alpha R^\lambda{}_{,\alpha} + D\xi^\lambda R^\alpha{}_{,\alpha} \\ & + Ej^\lambda R^\alpha{}_{,\alpha} + Fg^{\lambda\mu}\xi_\alpha R^\alpha{}_{,\mu} + Gg^{\lambda\mu}j_\alpha R^\alpha{}_{,\mu},\end{aligned}$$

where $A, B, C \cdots$ are functions of ξ^λ. We note that j^λ and ξ^λ are arbitrary. Hence, consider the transformation

$$(5.2)\qquad {}'j^\lambda = j^\lambda, \qquad {}'\xi^\lambda = k\xi^\lambda,$$

where k is an arbitrary constant. By substituting (5.2) into (5.1), we see that $\bar{R}^\lambda$ is composed of two independent vectors:

$$\begin{aligned}(5.3)\qquad \bar{R}^\lambda &= AR^\lambda + B\xi^\alpha R^\lambda{}_{,\alpha} + D\xi^\lambda R_{,\alpha} + Fg^{\lambda\mu}\xi_\alpha R^\alpha{}_{,\mu}, \\ {}'\bar{R}^\lambda &= Cj^{\alpha\prime}R^\lambda{}_{,\alpha} + Ej^{\lambda\prime}R^\alpha{}_{,\alpha} + Gg^{\lambda\mu}j_\alpha{}'R^\alpha{}_{,\mu}.\end{aligned}$$

Forming the divergence of the first vector in (5.3) we obtain

$$\begin{aligned}(5.4)\qquad \bar{R}^\lambda{}_{,\lambda} = {} & (A + B + 3D + F + \xi^\lambda D_{,\lambda})R^\alpha{}_{,\alpha} + (B + D)\xi^\alpha R^\lambda{}_{,\alpha\lambda} \\ & + Fg^{\lambda\mu}\xi_\alpha R^\alpha{}_{,\mu\lambda} + R^\lambda A_{,\lambda} + \cdots,\end{aligned}$$

where the dots denote terms containing the derivatives of B, F. Since R^λ is arbitrary, we may imagine R^λ expanded in a power series in which the derivatives $R^\lambda{}_{,\alpha}$, $R^\lambda{}_{,\alpha\beta}$ are arbitrary at any given point. Hence, the condition that the divergence of $\bar{R}^\lambda$ vanish for arbitrary R^λ, ξ^λ implies that

$$(5.5)\qquad F = 0, \qquad B + D = 0, \qquad A + B + 3D + F = 0,$$

and A, B, D, F are constants. These results also follow from the fact that

the identical vanishing of (5.4) implies the vanishing of the coefficients of all invariants of (5.4). From (5.5), we find

$$B = -D, \qquad A = -2D. \tag{5.6}$$

Thus, we may represent the most general vector of the first type, $\bar{R}^\lambda$, whose divergence is identically zero and which depends linearly upon R^λ and its derivatives, by

$$\bar{R}^\lambda = -2R^\lambda - \xi^\alpha R^\lambda{}_{,\alpha} + \xi^\lambda R^\alpha{}_{,\alpha}. \tag{5.7}$$

Similarly, one may show that the most general vector of the type $'\bar{R}^\lambda$ (see 5.3) may be represented by

$$'\bar{R}^\lambda = j^{\alpha'} R^\lambda{}_{,\alpha} - j^{\lambda'} R^\alpha{}_{,\alpha}. \tag{5.8}$$

By carrying out a similar analysis, one may show that the only possible second and third order tensors are

$$\bar{R}^{\mu\alpha} = -2R^{\mu\alpha} - \xi^\gamma R^{\mu\alpha}{}_{,\gamma} + \xi^\alpha R^{\mu\gamma}{}_{,\gamma}, \tag{5.9}$$

$$'\bar{R}^{\mu\alpha} = j^{\gamma'} R^{\mu\alpha}{}_{,\gamma} - j^{\alpha'} R^{\mu\gamma}{}_{,\gamma}, \tag{5.10}$$

$$\bar{R}^{\lambda\mu\alpha} = -2R^{\lambda\mu\alpha} - \xi^\gamma R^{\lambda\mu\alpha}{}_{,\gamma} + \xi^\alpha R^{\lambda\mu\gamma}{}_{,\gamma}, \tag{5.11}$$

$$'\bar{R}^{\lambda\mu\alpha} = j^{\gamma'} R^{\lambda\mu\alpha}{}_{,\gamma} - j^{\alpha'} R^{\lambda\mu\gamma}{}_{,\gamma}. \tag{5.12}$$

b. *The Equivalence of Our Construction of Tensors and the Above Construction.* It is evident that one could describe the basic tensors of the theory by equations (5.7) through (5.12). However, the construction used in section 4 is an equivalent construction, as will be shown. Further, this previous construction is easier to evaluate and shows why the barred tensors depend upon independent scalars. In order to show the equivalence of the barred tensors in sections 4 and 5, we carry through the computations of section 4.

Consider the equations (4.1), (4.4). By eliminating $\bar{S}_\gamma$, $'\bar{S}_\gamma$ we find

$$\bar{R}^\alpha = -\delta_{\lambda\beta}{}^{\alpha\rho}\partial(R^\lambda \xi^\beta)/\partial\xi^\rho, \qquad '\bar{R}^\alpha = -\delta_{\lambda\beta}{}^{\alpha\rho}\partial('R^\lambda j^\beta)/\partial\xi^\rho. \tag{5.13}$$

Replacing $\delta_{\lambda\beta}{}^{\alpha\rho}$ by $\delta_\lambda{}^\alpha \delta_\beta{}^\rho - \delta_\beta{}^\alpha \delta_\lambda{}^\rho$ and carrying through the differentiations in (5.13), we obtain the relations (5.7), (5.8). Similarly, one may show that (4.2), (4.5) and (4.3), (4.6) lead to (5.9), (5.10) and (5.11), (5.12), respectively.

c. *The Scalars of $\bar{R}^\lambda$ etc., are Linear Functions of the Scalars of R^λ etc., and Their Derivatives* (see section 4b, remark 1). In order to verify this result, we return to equation (3.7) and the equations (3.1), (3.2), (3.3)

defining the structure of R^λ, $R^{\lambda\mu}$, $R^{\lambda\mu\alpha}$. From these equations and the formulas (5.7), (5.8), etc., it is evident that the scalars of $\bar{R}^\lambda$ etc., are linear functions of the scalars of R^λ and their derivatives. A similar argument can be used if three independent vector fields ξ^λ, j^λ, i^λ exist.

d. *Do the $\bar{R}^\lambda$, etc., Have the Structure of Correlation Tensors?* The argument of (c) shows that the $\bar{R}^\lambda$ etc., do have the structure of correlation tensors. Note, it is essential that $\bar{R}^\lambda$ be a *linear homogeneous function in ξ^λ, j^λ, i^λ of order zero* (etc., for $\bar{R}^{\lambda\mu}$, $\bar{R}^{\lambda\mu\alpha}$, see section 4a, hypothesis 1).

6. The symmetry of $R^{\mu\alpha}$. As we have noted, the above procedure does not assure us of the symmetry of $\bar{R}^{\mu\alpha}$. When a single independent vector field ξ^λ exists, then $\bar{R}^{\mu\alpha}$ is symmetric. Hence, no problem exists for this case. For the case in which two independent vector fields ξ^λ, j^λ exist, the following tensor is symmetric (if $R^{\mu\alpha}$ is symmetric) and has vanishing divergence

$$(6.1) \qquad \bar{R}^{\mu\alpha} = j^\alpha j^\gamma R^{\mu\rho}{}_{,\gamma\rho} + j^\mu j^\gamma R^{\alpha\rho}{}_{,\gamma\rho} - j^\alpha j^\mu R^{\rho\gamma}{}_{,\gamma\rho} - j^\gamma j^\rho R^{\mu\alpha}{}_{,\gamma\rho}.$$

It should be noted that the tensor $\bar{R}^{\mu\alpha}$ of (6.1) is linear in the second derivatives of $R^{\mu\alpha}$. In fact, if we change the hypothesis (1) of section (4a) to read:

(1)′ *$\bar{R}^{\mu\alpha}$ is symmetric and is a linear function of $R^{\lambda\mu}$ and its first and second derivatives with respect to ξ^λ,*

then present computations indicate that (6.1) furnishes the only solution to the problem. If an additional independent vector field i^λ is present then an additional tensor $\bar{R}^{\lambda\mu}$ can be formed. The j^λ in (6.1) is merely replaced by i^λ.

UNIVERSITY OF MICHIGAN.

HOMOGENEOUS DIRICHLET PROBLEM FOR INHOMOGENEOUS ULTRAHYPERBOLIC EQUATION.*

By O. G. Owens.

1. Introduction. There is a class [1] of 4-dimensional domains G for which the solution of the Dirichlet problem for the ultrahyperbolic equation,

$$L[u] \equiv u_{x_1x_1} + u_{x_2x_2} - u_{y_1y_1} - u_{y_2y_2} = f(x_1, x_2, y_1, y_2), \tag{1.1}$$

is unique. Closely related uniqueness questions are discussed in a paper by F. John [2]. The existence of another such class of domains is established in the present paper. Without now specifying the differentiability assumptions on $u(x_1, x_2, y_1, y_2)$, this new class of domains can be described analytically as follows: Let $G(X)$ and $G(Y)$ be regions in X and Y space, respectively (see section 2), and let $\lambda(X)$ and $\mu(Y)$ denote the characteristic values of $G(X)$ and $G(Y)$. Then, if $|\lambda(X) - \mu(Y)|$ is never zero, the Dirichlet problem for (1.1) is unique for the cross product region $G = G(X) \times G(Y)$. Moreover, if f vanishes on G^*, and if $|\lambda(X) - \mu(Y)|$ has a positive lower bound, then a solution of (1.1) exists which vanishes on G^*. The fact that regions G exist for which $|\lambda(X) - \mu(Y)|$ is bounded away from zero is shown by means of an example.

2. Notation. $G(X)$ will denote a simply-connected open region in 2-dimensional cartesian (x_1, x_2)-space—more briefly X-space. $x \equiv (x_1, x_2)$ will always denote a point of $G(X)$, and $x^* \equiv (x^*_1, x^*_2)$ a point of the boundary $G^*(X)$ of $G(X)$. It will be assumed that $G^*(X)$ consists of a finite number of piecewise analytic arcs.

An arbitrary characteristic value and associated characteristic function of the region $G(X)$ will be denoted by $\lambda(X)$ and $\phi(X)$. That is, $\lambda(X)$ and $\phi(X)$ are solutions of the following equations:

$$\Delta_x \phi(x) + \lambda(X)\phi(x) = 0, \qquad x \in G(X); \tag{2.1}$$

$$\phi(x) = 0, \qquad x \in G^*(X); \qquad \int_{G(X)} \phi^2(x)\,dx = 1,$$

where $\Delta_x = \partial^2/\partial x_1^2 + \partial^2/\partial x_2^2$, $dx = dx_1 dx_2$. As in the case of $G(X)$, a

* Received October 16, 1950; revised July 31, 1951.

region $G(Y)$ is defined in 2-dimensional cartesian Y-space, together with the related characteristic values $\mu(Y)$ and associated characteristic functions $\psi(y)$. The 4-dimensional cartesian cross product region consisting of all points $(x, y) = ((x_1, x_2, y_1, y_2)$ such that $x \varepsilon G(X)$ and $y \varepsilon G(Y)$ is denoted by $G = G(X) \times G(Y)$, and its boundary by G^*.

3. Uniqueness. Our proof of the uniqueness theorem involves the application of two lemmas having to do with uniform convergence of series of characteristic functions. It is convenient to introduce some more notation for the statement of these lemmas. We write $u = u(x, y) \varepsilon C^k(X, Y)$, if u has all partial derivatives of order k, and these are continuous functions of (x, y) on $\bar{G}$, the closure of G. We write $u = u(x, y) \varepsilon C^k(X)\,(C^k(Y))$, if u has all partial derivatives with respect to $x(y)$ of order k, and these are continuous functions of (x, y) on $\bar{G}$.

Henceforth in this paper we assume that

$$\begin{aligned} &u(x^*, y) = 0, \qquad (x^*, y) \varepsilon G^*, x^* \varepsilon G^*(X); \\ &u(x, y^*) = 0, \qquad (x, y^*) \varepsilon G^*, y^* \varepsilon G^*(Y). \end{aligned} \tag{3.1}$$

The first of our two lemmas is the following:

LEMMA 1. *If $u(x, y) \varepsilon C^4(X, Y)$, then $u(x, y)$ is equal to the uniformly, absolutely convergent series*

$$u(x, y) = \sum_{n,m=1}^{\infty} a_{nm}\phi_n(x)\psi_m(y), \qquad a_{nm} = \int_{G(X)} dx\phi_n(x) \int^{G(Y)} u(x, y)\psi_m(y)dy. \tag{3.2}$$

Note. As proved, Lemma 1 is true under weaker differentiability restrictions on $u(x, y)$, namely, u, $\Delta_x u$, $\Delta_y u \varepsilon C^2(X)$ and $C^2(Y)$.

Proof. As $u(x, y) \varepsilon C^2(X)$ and satisfies (3.1), it has the series representation [3]:

$$u(x, y) = \sum_{n=1}^{\infty} \phi_n(x) \int_{G(X)} u(x, y)\phi_n(x)dx. \tag{3.4}$$

As $\int_{G(X)} u(x, y)\phi_n(x)dx \varepsilon C^2(Y)$ and vanishes on $G^*(Y)$, (3.4) can be expressed as follows:

$$u(x, y) = \sum_{n=1}^{\infty} \phi_n(x) \sum_{m=1}^{\infty} \psi_m(y) \int_G u(x, y)\phi_n(x)\psi_m(y)dxdy. \tag{3.5}$$

Because of (2.1) and (3.1), the integrals in (3.5) can be so transformed by Green's theorem that $u(x, y)$ can now be expressed as follows:

$$(3.6) \qquad u(x,y) = \sum_{n=1}^{\infty} \phi_n(x) \sum_{m=1}^{\infty} \psi_m(y) \int_G [\Delta_y \Delta_x u/(\lambda_n \lambda_m)] \phi_n(x) \psi_m(y) dx dy.$$

We mention that the above use of Green's theorem is all right, as the first derivatives of $\phi_n(x)$ $(\psi_m(y))$ are uniformly bounded [4] on $G(X)(G(Y))$.

The expansion (3.2) will be justified by establishing the uniform absolute convergence of the series (3.6). Our proof will be based on the inequalities of Schwarz and Bessel, and is as follows:

(3.7)

$$|u(x,y)| \leq \sum_{n=1}^{\infty} |\phi_n(x)/\lambda_n| \sum_{m=1}^{\infty} |\psi_m(y)/\mu_m| \cdot \left| \int_{G(Y)} dy \psi_m(y) \int_{G(X)} [\Delta_y \Delta_x u] \phi_n(x) dx \right|$$

$$\leq \sum_{n=1}^{\infty} |\phi_n(x)/\lambda_n| \cdot \{ [\sum_{m=1}^{\infty} \psi_m^2(y)/\mu_m^2] \cdot \int_{G(Y)} dy [\int_{G(X)} [\Delta_y \Delta_x u] \phi_n(x) dx]^2 \}^{\frac{1}{2}}$$

$$\leq [\sum_{n=1}^{\infty} \phi_n^2(x)/\lambda_n^2]^{\frac{1}{2}} \cdot [\sum_{m=1}^{\infty} \psi_m^2(y)/\mu_m^2]^{\frac{1}{2}} \cdot [\int_{G(Y)} dy \sum_{n=1}^{\infty} \{ \int_{G(X)} [\Delta_y \Delta_x u] \phi_n(x) dx \}^2]^{\frac{1}{2}}$$

$$\leq [\sum_{n=1}^{\infty} \phi_n^2(x)/\lambda_n^2]^{\frac{1}{2}} \cdot [\sum_{m=1}^{\infty} \psi_m^2(y)/\mu_m^2]^{\frac{1}{2}} \cdot [\int_{G(Y)} dy \int_{G(X)} [\Delta_y \Delta_x x]^2 dx]^{\frac{1}{2}}.$$

Hence, as $\Sigma \phi_n^2(x)/\lambda_n^2$ and $\Sigma \psi_m^2(y)/\mu_m^2$ are uniformly bounded [5], Lemma 1 is now proved.

The last lemma we need is the following:

LEMMA 2. *If* $u(x, y) \in C^6(X, Y)$ *and* $\Delta_x u = \Delta_y u$, *then*

$$(3.8) \qquad H(x,y) \equiv \sum_{n,m=1}^{\infty} \lambda_n a_{nm} \phi_n(x) \psi_m(y) = -\Delta_x u,$$

$$(3.9) \qquad I(x,y) \equiv \sum_{n,m=1}^{\infty} \mu_m a_{nm} \phi_n(x) \psi_m(y) = -\Delta_y u.$$

Note. As proved, Lemma 2 is true if $u(x, y)$ is such that $u(x, y) \in C^2(X$, and $C^2(Y)$, $\Delta_x u \in C^2(X)$, $\Delta_x{}^2 u \in C^2(Y)$, $\Delta_y u \in C^2(Y)$, $\Delta_y{}^2 u \in C^2(X)$.

Proof. We first show that the series (3.8) is uniformly absolute convergent. Because of (2.1) and (3.3),

$$(3.10) \qquad a_{nm} = -\lambda_n^{-2} \int_{G(Y)} dy \psi_m(y) \int_{G(X)} [\Delta_x u] \phi_n(x) dx.$$

As $\Delta_x u(x, y)_{x=x^*} = \Delta_y u(x^*, y) = 0$, (3.10) can be transformed as follows:

$$(3.11)\qquad a_{nm} = \lambda_n^{-2} \int_{G(Y)} dy \psi_m(y) \int_{G(X)} \Delta_x{}^2 u \phi_n(x)\, dx$$

$$a_{nm} = (-\mu_m \lambda_n{}^2)^{-1} \int_{G(X)} dx \phi_n(x) \int_{G(Y)} [\Delta_y \Delta_x{}^2 u] \psi_m(y)\, dy.$$

The last integral representation being valid since $\Delta_x{}^2 u(x, y)_{y=y^*} = 0$.

$H(x, y)$ can now be majorized as follows:

(3.12)

$$|H(x,y)| \leq \sum_{m=1}^{\infty} |\psi_m(y)/\mu_m| \sum_{n=1}^{\infty} |\phi_n(x)/\lambda_n| \cdot \left| \int_{G(X)} dx \phi_n(x) \int_{G(Y)} [\Delta_y \Delta_x{}^2 u] \psi_m(y) dy \right|$$

$$\leq \sum_{m=1}^{\infty} |\psi_m(y)/\mu_m| \cdot \{[\sum_{n=1}^{\infty} \phi_n{}^2(x)/\lambda_n{}^2] \cdot [\int_{G(X)} dx \{ \int_{G(Y)} [\Delta_y \Delta_x{}^2 u] \psi_m(y) dy \}^2]\}^{\frac{1}{2}}$$

$$\leq [\sum_{n=1}^{\infty} \phi_n{}^2(x)/\lambda_n{}^2]^{\frac{1}{2}} \cdot [\sum_{m=1}^{\infty} \psi_m{}^2(y)/\mu_m{}^2]^{\frac{1}{2}} \cdot [\int_{G(X)} dx \int_{G(Y)} [\Delta_y \Delta_x{}^2 u]^2 dy]^{\frac{1}{2}}.$$

The above inequalities are obtained, as in (3.7), by repeated applications of the inequalities of Schwarz and Bessel. The inequality (3.12) implies that the series (3.8) has the desired convergence property.

We shall now show that $H(x, y) = -\Delta_x u(x, y)$. Let $K \equiv K(x, z)$ denote the Green's function for the region $G(X)$, and recall that the characteristic value and function are solutions of the equation

$$\phi_n(x) = \lambda_n \int_{G(X)} K(x, z) \phi_n(z)\, dz.$$

Thus, because of the uniform convergence of the series $H(x, y)$,

$$(3.13)\qquad \int_{G(X)} K(x, z) H(z, y)\, dz = \sum_{n,m=1}^{\infty} a_{nm} \psi_m(y) \lambda_n \int_{G(X)} K(x, z) \phi_n(z)\, dz$$

$$= \sum_{n,m=1}^{\infty} a_{nm} \psi_m(y) \phi_n(x) = u(x, y).$$

Moreover, as $u \in C^2(X)$, $u(x, y)$ can be represented in the form

$$(3.14)\qquad -\int_{G(X)} K(x, z) \Delta_x u(z, y)\, dz = u(x, y).$$

Consequently, because of (3.13) and (3.14), we have the following identity:

$$(3.15)\qquad V(x, y) \equiv \int_{G(X)} K(x, z) [H(z, y) + \Delta_x u(x, y)_{x=z}] dz \equiv 0.$$

Let $g(z, y)$ denote the bracketed expression in the integrand of (3.15). Then,

$$(3.16)\qquad V(x, y) \equiv \int_{G(X)} K(x, z) g(z, y)\, dz \equiv 0.$$

Our proof that $g \equiv 0$ will follow the usual procedure of potential theory. Let Ω denote an arbitrary area with closure in $G(X)$, and Ω^* the boundary of Ω. Represent the normal derivative with respect to Ω^* by d/dn_{Ω^*}. Then,

$$0 = \int_{\Omega^*} dV/dn_{\Omega^*} d\Omega^*$$

$$= \int_{\Omega^*} d\Omega^* \int_{G(X)} dK/dn_{\Omega^*} g(z,y)\,dz = \int_{G(X)} dz g(z,y) \int_{\Omega^*} dK/dn_{\Omega^*} d\Omega^*.$$

The interchange of the order of integration being permissible because of the summability of the normal derivative of K [4]. Now, as is well-known, the last (interior) integral is -2π if $z \subset \Omega$, and 0 if $z \subset G(x) - \Omega$. Therefore,

$$-2\pi \int_{\Omega} g(z,y)\,dz = 0.$$

As Ω has been chosen arbitrarily, and as g is continuous, we conclude that $g(x,y) \equiv 0$, that is, $-\Delta_x u(x,y) \equiv H(x,y)$. As the relation $-\Delta_y u(x,y) \equiv I(x,y)$ follows in the same way, Lemma 2 is proved.

We now consider the proof of the uniqueness result, which is stated as follows:

THEOREM 1. *Assume $u(x,y)$ is of class $C^6(X,Y)$, is a solution of the equation $\Delta_x u = \Delta_y u$, $(x,y) \in G \equiv G(X) \times G(Y)$ and assumes vanishing boundary values on G^*. Then, provided $|\lambda(X) - \mu(Y)| \neq 0$, we must have $u(x,y) \equiv 0$ on G.*

Note. Under the following weaker differentiability requirements on u, Theorem 1 is still true:

$$(3.17) \qquad u, \Delta_x u, \Delta_y u \in C^2(X) \text{ and } C^2(Y);\ \Delta_x{}^2 u \in C^2(Y),\ \Delta_y{}^2 u \in C^2(X).$$

In general, functions possessing the differentiability properties (3.17) on G are said to be of class D.

Proof. By Lemma 1,

$$(3.18) \qquad u(x,y) = \sum_{n,m=1}^{\infty} a_{nm} \phi_n(x) \psi_m(y),$$

and by Lemma 2,

$$(3.19) \qquad 0 = \Delta_x u(x,y) - \Delta_y u(x,y) = \sum_{n,m=1}^{\infty} a_{nm}(\mu_m - \lambda_n) \phi_n(x) \psi_m(y).$$

Thus, as (3.19) is uniformly absolutely convergent, and as the system $\{\phi_n(x)\psi_m(y)\}$ $(n, m = 1, 2, 3, \cdots)$ is complete on G, $a_{nm}(\mu_m - \lambda_n) = 0$.

Hence, as the characteristic values λ_n and μ_m are distinct by assumption, $a_{nm} = 0$ $(n, m = 1, 2, 3, \cdots)$, that is, $u(x, y) = 0$ for all $(x, y) \varepsilon G$.

4. Existence. The existence theorem which we establish has to do with a homogeneous boundary value problem—boundary values vanishing on G^*—for the non-homogeneous equation (1.1). The proof presupposes the vanishing of f on G^*. The restriction that $G = G(X) \times G(Y)$ is such that $|\lambda(X) - \mu(Y)|$ is uniformly bounded away from zero, is an indispensable requirement in the establishment of the existence theorem. Domains with this property are exhibited in the next section.

Before stating our theorem we introduce some new symbols.

$$D_x^k f(x, y)\,(D_y^k f(x, y))$$

will denote any partial derivative of $f(x, y)$ with respect to $x(y)$ of order k; $D^k f(x, y)$ will denote any partial derivative of $f(x, y)$ of order k. $(D^0 f(x, y) = f(x, y))$.

THEOREM 2. *Assume $G = G(X) \times G(Y)$ is such that $|\lambda(X) - \mu(Y)|$ is bounded away from zero, and that $f(x, y) \varepsilon C^{24}(X, Y)$ on G. Moreover, assume that $D^k f(x, y)$, $0 \leq k \leq 22$, vanishes on G^*. Then there is a $u(x, y)$ of class D (see (3.17)) which vanishes on G^*, and is a solution of (1.1) on G.*

Proof. Assuming the existence of such a $u(x, y)$, we first determine its series expansion. If $u \varepsilon D$ and is a solution of (1.1), then the results of Lemma 2 are still valid, provided, as we have supposed, that f vanishes on G^*. Thus

$$u(x, y) = \sum_{n,m=1}^{\infty} a_{nm}\phi_n(x)\psi_m(y), \tag{4.1}$$

$$\Delta_x u - \Delta_y u = \sum_{n,m=1}^{\infty} a_{nm}(\mu_m - \lambda_n)\phi_n(x)\psi_m(y). \tag{4.2}$$

Now, as $f \varepsilon D$, Lemma 1 yields

$$f(x, y) = \sum_{n,m=1}^{\infty} b_{nm}\phi_n(x)\psi_m(y), \tag{4.3}$$

$$b_{nm} = \int_{G(X)} dx\phi_n(x) \int_{G(Y)} f(x, y)\psi_m(y)\,dy. \tag{4.4}$$

Consequently, as $\{\phi_n(x)\psi_m(y)\}$ is complete on G and as (1.1) is valid, we conclude that

$$a_{nm}(\mu_m - \lambda_n) = b_{nm} \qquad (n, m = 1, 2, 3, \cdots). \tag{4.5}$$

Thus, as $(\mu_m - \lambda_n)$ is never zero,

$$u(x,y) = \sum_{n,m=1}^{\infty} b_{nm}/(\mu_m - \lambda_n)\phi_n(x)\psi_m(y). \tag{4.6}$$

Our immediate purpose is to show that the restrictions imposed upon $f(x,y)$ in the statement of the theorem are sufficient to assure that $u \in D$, and that $\Delta_x u - \Delta_y u = f(x,y)$. Such a $u(x,y)$ is unique by Theorem 1.

The essential requirements for such a study are bounds on $|D^k\phi_n(x)|$ $(0 \leq k \leq 2)$. Such bounds are readily found in the literature for domains $G(X)$ having simple closed analytic curves as boundary. Moreover, such bounds are effective in the case of our example, section 5.

In the case where $G^*(X)$ consists of a finite number of analytic arcs, A. Hammerstein [4] has found the following bounds for the characteristic functions $\phi_n(x)$:

$$|\phi_n(x)| < A\lambda_n \quad (n = 1, 2, 3, \cdots); \tag{4.7}$$

$$|D'\phi_n(x)| < A\lambda_n^2. \tag{4.8}$$

If $G^*(X)$ consists of a simple closed analytic curve, J. Schnauder [6] has determined bounds for the second and third partial derivatives of the solution of a general elliptic equation in terms of bounds for the solution and its first partial derivatives. The general equation considered by Schauder includes as a special case the Poisson equation: $\partial^2 v/\partial x_1^2 + \partial^2 v/\partial x_2^2 = \rho(x_1, x_2)$. Consequently, because of (4.7) and (4.8), Schauder's results enable us to conclude immediately that

$$|D^2\phi_n(x)| < A\lambda_n^3; \tag{4.9}$$

$$|D^3\phi_n(x)| < A\lambda_n^4. \tag{4.10}$$

For the special domain $G(X)$ considered in the next section, the characteristic functions $\phi_n(x)$ are given explicitly, and the fact that they satisfy the above bounds is apparent.

It is now possible to show that $u \in D$, and we commence this investigation. In a purely formal manner, (4.6) yields the equations

$$\begin{aligned} D_x^2 u(x,y) &= \sum_{n,m=1}^{\infty} b_{nm}/(\mu_m - \lambda_n)\psi_m(y) D_x^2\phi_n(x), \\ D_y^2 u(x,y) &= \sum_{n,m=1}^{\infty} b_{nm}/(\mu_m - \lambda_n)\phi_n(x) D_y^2\psi_m(y). \end{aligned} \tag{4.11}$$

Therefore, because of (4.10), if

$$\sum_{n,m=1}^{\infty} |b_{nm}|\,\mu_m\lambda_n^3 < \infty, \quad \text{and} \quad \sum_{n,m=1}^{\infty} |b_{nm}|\,\lambda_n\mu_m^3 < \infty, \tag{4.12}$$

then $u(x, y)$ exists and $\varepsilon C^2(X)$ and $C^2(Y)$. Consequences of (4. 12) are the next two equations,

$$\Delta_x u(x, y) = -\sum_{n,m=1}^{\infty} \lambda_n b_{nm}/(\mu_m - \lambda_n)\phi_n(x)\psi_m(y), \tag{4. 13}$$

$$\Delta_y u(x, y) = -\sum_{n,m=1}^{\infty} \mu_m b_{nm}/(\mu_m - \lambda_n)\phi_n(x)\psi_m(y), \tag{4. 14}$$

which show that

$$\Delta_x u - \Delta_y u = \sum_{n,m=1}^{\infty} b_{nm}\phi_n(x)\psi_m(y) = f(x, y). \tag{4. 15}$$

From (4. 13) and (4. 14), we conclude that both $\Delta_x u$ and $\Delta_y u$ are elements of $C^2(X)$ and $C^2(Y)$, if

$$\begin{aligned} &\sum_{n,m=1}^{\infty} | b_{nm} | \mu_m \lambda_n{}^4 < \infty, \qquad \sum_{n,m=1}^{\infty} | b_{nm} | \lambda_n{}^2 \mu_m{}^3 < \infty, \\ &\sum_{n,m=1}^{\infty} | b_{nm} | \lambda_n \mu_m{}^4 < \infty, \qquad \sum_{n,m=1}^{\infty} | b_{nm} | \mu_m{}^2 \lambda_n{}^3 < \infty. \end{aligned} \tag{4. 16}$$

The inequalities (4. 16) imply that

$$\Delta_x{}^2 u = \sum_{n,m=1}^{\infty} \lambda_n{}^2 b_{nm}/(\mu_m - \lambda_n)\phi_n(x)\psi_m(y), \tag{4. 17}$$

$$\Delta_y{}^2 u = \sum_{n,m=1}^{\infty} \mu_m{}^2 b_{nm}/(\mu_m - \lambda_n)\phi_n(x)\psi_m(y). \tag{4. 18}$$

Hence, $\Delta_x{}^2 u \, \varepsilon \, C^2(Y)$, and $\Delta_y{}^2 u \, \varepsilon \, C^2(X)$, if

$$\sum_{n,m=1}^{\infty} | b_{nm} | \mu_m{}^3 \lambda_3{}^3 < \infty. \tag{4. 19}$$

Therefore, it will be sufficient to show that

$$\sum_{n,m=1}^{\infty} | b_{nm} | \mu_m{}^4 \lambda_m{}^4 < \infty, \tag{4. 20}$$

in order to conclude that $u(x, y) \, \varepsilon \, D$. As $\Sigma \, 1/\lambda_n{}^2 < \infty$ and $\Sigma \, 1/\mu_m{}^2 < \infty$, the inequality (4. 20) is certainly true if

$$| b_{nm} | < B/(\lambda_n{}^6 \mu_m{}^6) \qquad (n, m = 1, 2, 3, \cdots). \tag{4. 21}$$

From (4. 4) and (2. 1) we deduce that

(4.22)

$$b_{nm} = \int_{G(Y)} dy\psi_m(y) \int_{G(X)} f(x, y)\phi_n(x)dx$$

$$= -1/\lambda_n \int_{G(Y)} dy\psi_m(y) \int_{G(X)} f(x, y)\Delta_x\phi_n(x)\,dx$$

$$= (-1/\lambda_n) \int_{G(Y)} dy\psi_m(y) \int_{G(X)} [\Delta_x{}^1 f]\phi_n(x)dx$$

$$= (-1/\lambda_n)^6 \int_{G(Y)} dy\psi_m(y) \int_{G(X)} [\Delta_x{}^6 f]\phi_n(x)dx.$$

In turn, (4.22) implies

$$b_{nm} = (\lambda_n\mu_m)^{-6} \int_{G(X)} dx\phi_n(x) \int_{G(Y)} [\Delta_y{}^6(\Delta_x{}^6 f)]\psi_m(y)dy. \tag{4.23}$$

An application of Schwarz's inequality to (4.23) yields the desired bound (4.21), namely,

$$|b_{nm}| \leq (\lambda_n\mu_m)^{-6}\Big[\int_G [\Delta_y{}^6(\Delta_x{}^6 f)]^2 dxdy\Big]^{\frac{1}{2}}.$$

This result completes the proof of Theorem 2.

5. An explicit domain G. We now propose to give an example of a region $G = G(X)\times G(Y)$ for which $|\lambda(X) - \mu(Y)|$ is bounded away from zero.

Take $G(X)$ and $G(Y)$ to be squares with sides of length a and b, respectively. More precisely, let

$$G(X) \equiv [0 \leq x_1, x_2 \leq a], \qquad G(Y) \equiv [0 \leq y_1, y_2 \leq b]. \tag{5.1}$$

Then, as is well-known, the corresponding characteristic values and functions are:

$$\phi_n(x) \equiv (2/a)\sin(n_1\pi/ax_1)\sin(n_2\pi/ax_2), \quad \lambda(X) \equiv \pi^2(n_1{}^2 + n_2{}^2)/a^2,$$
$$\psi_m(y) \equiv (2/b)\sin(n_3\pi/by_1)\sin(n_4\pi/by_2), \quad \mu(Y) \equiv \pi^2(n_3{}^2 + n_4{}^2)/b^2. \tag{5.2}$$

As suggested by Mr. Louis Nirenberg and Mr. George Raney, we take

$$a = 3^{-\frac{1}{2}}R \quad \text{and} \quad b = R, \tag{5.3}$$

R being any real positive number. Then

$$|\lambda(X) - \mu(Y)| = (\pi^2/R^2)\,|\,3(n_1{}^2 + n_2{}^2) - (n_3{}^2 + n_4{}^2)\,|.$$

We shall now prove that

$$(5.4) \qquad |3(n_1^2+n_2^2)-(n_3^2+n_4^2)| \qquad (n_1, n_2, n_3, n_4 = 1, 2, 3, \cdots)$$

is never zero. A consequence of this will be $|\lambda(X)-\mu(Y)| \geq \pi^2/R^2$, a positive lower bound.

Suppose there exists a non-vacuous set of points (m_1, m_2, m_3, m_4) for which (5.4) is zero. There is then a point (m_1, m_2, m_3, m_4) of this set for which $m_3^2+m_4^2$ is a minimum. Thus, as any square is congruent either to zero or one modulo three, there are integers p_3 and p_4 such that $m_3^2 = 3^2p_3^2$ and $m_4^2 = 3^2p_4^2$. Hence, $m_1^2+m_2^2 = 3(p_3^2+p_4^2)$. The same reasoning shows the existence of integers p_1 and p_2 such that $m_1^2 = 3^2p_1^2$ and $m_2^2 = 3^2p_2^2$. Consequently, $3(p_1^2+p_2^2) = (p_3^2+p_4^2) < (m_3^2+m_4^2)$, which is a contradiction.

UNIVERSITY OF WISCONSIN.

REFERENCES.

[1] O. G. Owens, "Uniqueness of solutions of ultrahyperbolic partial differential equations," *American Journal of Mathematics*, vol. 69 (1947), pp. 184-188; especially p. 185.

[2] F. John, "On linear partial differential equations with analytic coefficients," *Communications on Pure and Applied Mathematics*, vol. II (1949), pp. 209-253.

[3] R. Courant and D. Hilbert, *Methoden der Mathematischen Physik*, vol. I (1931), p. 319.

[4] A. Hammerstein, "Über Entwicklungen gegebener Funktionen nach Eigenfunktionen von Randwertaufgaben," *Mathematische Zeitschrift*, vol. 27 (1927), pp. 269-311; especially pp. 290-294.

[5] R. Courant and D. Hilbert, *Methoden der Mathematische Physik*, vol. I (1931), p. 114.

[6] J. Schauder, "Über den Zusammenhang zwischen der Eindeutigkeit und Lösbarkeit partieller Differentialgleichungen zweiter Ordnung vom elliptischen Typus," *Mathematische Annalen*, vol. 106 (1932), pp. 661-721; especially pp. 687-688.

OSCILLATION THEOREMS.*

By CHOY-TAK TAAM.

1. Introduction. In the present paper we shall study the distribution of the complex zeros of the analytic solutions $W(z)$ of

$$d^2W/dz^2 + Q(z)W = 0, \tag{1.1}$$

where $Q(z)$ is an analytic function of the complex variable z in the domain concerned.

The method we employ in this paper is to derive a pair of simultaneous differential equations (2.7) and (2.8) which are satisfied by the modulus and argument of any solution $W(z)$ along the real axis. Since the modulus of a solution and the solution itself have exactly the same zeros, the equations (2.7) and (2.8) reflect the distribution of the zeros of the solutions on the real axis. The results so obtained can be extended to the zeros of the solutions along any analytic curve, in particular. the linear segments in the z-plane.

Using equations (2.7) and (2.8), methods will be developed to determine regions in which there are no zeros and thus we may confine them to a certain smaller known region of the z-plane. These methods could to some extent complete the methods based on the Green transform as developed by E. Hille.[1]

2. The associated differential equations. On the real axis, equation (1.1) has the form

$$d^2W/dx^2 + Q(x) = 0, \tag{2.1}$$

where x is the real part of the complex variable z. Let $Q(z)$ be analytic along the real axis. In the sequel, by a solution $W(x)$ of (2.1) is always meant an analytic solution, not identically zero, unless otherwise stated.

Let

$$Q(x) = q_1 + iq_2, \qquad W(x) = u + iv, \qquad W(x) = re^{i\theta}, \tag{2.2}$$

* Received January 15, 1951; revised April 6, 1951.

The author is grateful to Professor E. Hille for some suggestions in the present paper.

[1] See E. Hille, "Oscillation theorems in the complex domain," *Transactions of the American Mathematical Society*, vol. 23 (1922), pp. 350-385.

where $q_1(x)$, $q_2(x)$, $u(x)$, $v(x)$, $r(x)$, $\theta(x)$ are real and $r(x) \geqq 0$. The derivatives $W'(x)$, $W''(x)$ with respect to x at any point where $W(x) \neq 0$ are

$$(2.3) \quad W' = (r' + ir\theta')e^{i\theta}, \qquad W'' = [\{r'' - r(\theta')^2\} + i\{r\theta'' + 2r'\theta'\}]e^{i\theta}.$$

Substituting the expressions (2.2) and (2.3) in the equation (2.1) and separating the real and imaginary parts gives

$$(2.4) \quad r'' + (q_1(x) - (\theta')^2)r = 0; \qquad (2.5) \quad r\theta'' + 2r'\theta' + q_2(x)r = 0.$$

Multiplying equation (2.5) by $r(x)$, it can be rewritten in the form

$$(2.6) \qquad (\theta' r^2)' + q_2(x)r^2 = 0.$$

The equations (2.4) and (2.6) hold at every point of the real axis where $W(x) \neq 0$. At a point where $W(x) = 0$, $r'(x)$, $r''(x)$, $\theta'(x)$ and $(\theta'(x)r^2(x))'$ do not exist. We shall, however, replace $r(x)$ and $\theta'(x)$ by functions $Y(x)$ and $\Phi(x)$ respectively (to be defined shortly) such that equations (2.4) and (2.6) will hold at every point of the real axis.

The new function $Y(x)$, called the *associated solution* of $W(x)$, is defined as follows. If $W(x)$ does not vanish on the real axis, let $Y(x) = r(x) = (u^2(x) + v^2(x))^{\frac{1}{2}}$, $-\infty < x < \infty$. If $W(x)$ vanishes at the points a_i, $a_i < a_{i+1}$, define $Y(x) = r(x)$ in $a_{2k} \leqq x \leqq a_{2k+1}$ and $Y(x) = -r(x)$ in $a_{2k-1} < x < a_{2k}$, where $k = 0, \pm 1, \pm 2, \cdots$.

The associated solution $Y(x)$ and $r(x)$ and $W(x)$ have exactly the same zeros. $Y(x)$ is obviously continuous for all x, and has continuous first and second derivatives at every point where $W(x)$ does not vanish. At every such point $Y(x)$ satisfies both equations (2.4) and (2.6).

We shall prove that $Y'(x)$ and $Y''(x)$ exist and are continuous at every zero a_i of $W(x)$. Let us consider the zero a_0. Since $W'(a_0) \neq 0$, from the definition of $Y(x)$ and the Law of the Mean, we have, as $x \to a_0 +$,

$$\begin{aligned} Y'(a_0 +) &= \lim [u(x)u'(x) + v(x)v'(x)][u^2(x) + v^2(x)]^{-\frac{1}{2}} \\ &= \lim [u'(c_1)u'(x) + v'(c_2)v'(x)][(u'(c_1))^2 + (v'(c_2))^2]^{-\frac{1}{2}} \end{aligned}$$

$= [(u'(a_0))^2 + (v'(a_0))^2]^{\frac{1}{2}} = |w'(a_0)|$, where $a_0 < c_1$, $c_2 < x$. Similarly $Y'(a_0 -) = |w'(a_0)|$. But $Y'(a_0) = \lim [Y(x) - Y(a_0)](x - a)^{-1}$, hence $Y'(a_0) = \lim Y'(c)$ where $c \to a_0$ and $a_0 < c < x$ or $a_0 > c > x$. Thus $Y'(x)$ exists and is continuous at a_0.

We shall soon see that $\theta'(x)$ has a limit at the zeros a_i of $W(x)$. In equation (2.4) replacing $r(x)$ by $Y(x)$, we find that $\lim Y''(x) = 0$ as $x \to a_0$. By the law of the mean,

$$Y''(a_0) = \lim_{x \to a_0} [Y'(x) - Y'(a_0)](x - a)^{-1} = \lim_{c \to a_0} Y''(c)$$

$$(a_0 < c < x \text{ or } a_0 > c > x).$$

Hence $Y''(x)$ exists and is continuous at a_0.

Between any two consecutive zeros, $a_i < x < a_{i+1}$, $W(x)$ can be represented by a continuous curve not passing through the origin in the W-plane. Then $\theta(x)$ can be so determined that it becomes a continuous function of x with continuous first derivative in $a_i < x < a_{i+1}$; namely,

$$\theta'(x) = [u(x)v'(x) - v(x)u'(x)][u^2(x) + v^2(x)]^{-1} \text{ if } a_i < x < a_{i+1}.$$

By de L'Hospital's rule it is easy to prove that

$$\lim \theta'(x) = 2^{-1}[u''(a_i)v'(a_1) - v''(a_i)u'(a_i)][(u'(a_i))^2 + (v'(a_i))^2]^{-1} = 0,$$

where $x \to a_i$. Define $\Phi(x) = \theta'(x)$ if $x \neq a_i$, and $\Phi(x) = 0$ if $x = a_i$, where $i = 0, \pm 1, \pm 2, \cdots$.

Clearly $\Phi(x)$ is continuous for all x, and from equation (2.6) it can be shown that $(\Phi(x)Y^2(x))'$ is also continuous for all x. Hence for each given solution $W(x)$ of (2.1), its associated functions $Y(x)$ and $\Phi(x)$ satisfy the simultaneous differential equations

$$(2.7) \qquad Y'' + (q_1 - \Phi^2)Y = 0; \qquad (2.8) \qquad (\Phi Y^2)' + q_2 Y^2 = 0.$$

3. A physical interpretation of equaltions (2.4) and (2.6). Consider a particle of unit mass whose position in a plane at time x is (r, θ), which moves under a force $q_1(x)r$ directed toward the origin and a force $q_2(x)r$ perpendicular to the radius vector but in the negative direction of θ. Equation (2.4) gives the relation of the forces acting along the radius vector, and equation (2.6) expresses the connection of the rate of change of the angular momentum to the moment of force.

4. Distribution of zeros on the real axis. In this section we study the distribution of the zeros of the solutions of (2.1) on the real axis by means of the equations (2.7) and (2.8). As in section 2, x is here a real variable.

THEOREM 4.1. *If $W(x)$ is any solution of (2.1), $q_2(x) \not\equiv 0$, and if $y(x)$ is a real-valued solution of $y'' + q_1(x)y = 0$, $y(x) \not\equiv 0$, finally if the zeros of $y(x)$ are b_i, $b_i < b_{i+1}$, $(i = 0, \pm 1, \pm 2, \cdots)$, then $W(x)$ has at most one zero in $b_i \leqq x \leqq b_{i+1}$.*

Proof. Suppose that $W(x)$ has more than one zero, and that one of them is a. From (2.8),

$$\Phi(x)Y^2(x) = -\int_a^x q_2(x)Y^2(x)\,dx,$$

because $W(x)$ and its associated solution $Y(x)$ have the same zeros. Since $q_2(x) \not\equiv 0$, $Y(x) \not\equiv 0$, then $\Phi(x) \not\equiv 0$. Now consider the equations

$$(4.1)\quad Y'' + (q_1(x) - \Phi(x))Y = 0; \qquad (4.2)\quad y'' + q_1(x)y = 0,$$

which are satisfied by $Y(x)$ and $y(x)$ respectively. By Sturm's comparison theorem,[2] we find that between any two consecutive zeros of $Y(x)$, there is at least one zero of $y(x)$. If $Y(x)$ has more than one zero in $b_i \leqq x \leqq b_{i+1}$, then $y(x)$ has at least one zero in $b_i < x < b_{i+1}$. Since $y(x)$ has no zero in $b_i < x < b_{i+1}$. then $Y(x)$ and consequently $W(x)$, has at most one zero in $b_i \leqq x \leqq b_{i+1}$.

THEOREM 4.2. *If $W(x)$ is a solution of (2.1), and if $q_2(x)$ changes sign at c_i, $c_i < c_{i+1}$ $(i = 0, \pm 1, \pm 2, \cdots)$, then $W(x)W'(x)$ has at most one zero in $c_i \leqq x \leqq c_{i+1}$.*

Proof. Suppose that $W(x)W'(x)$ has at least two zeros a_1 and a_2, $a_1 < a_2$. Since $W(a_i) = 0$ implies $Y(a_i) = 0$ and $W'(a_i) = 0$ implies $\Phi(a_i) = 0$, integrating equation (2.8) gives $\int_{a_1}^{a_2} q_2(x)Y^2(x)\,dx = 0$. Hence $q_2(x)$ must change sign at least once in $a_1 < x < a_2$. If $W(x)W'(x)$ has two zeros in $c_i \leqq x \leqq c_{i+1}$, then $q_2(x)$ must change sign at least once in $c_i < x < c_{i+1}$. Since $q_2(x)$ does not change sign in $c_i < x < c_{i+1}$, then $W(x)W'(x)$ has at most one zero in $c_i \leqq x \leqq c_{i+1}$.

5. Zero-free intervals. Suppose that $W(x)$ has a known zero, and without loss of generality, it is assumed to be at the origin. The theorems in this section will provide methods to determine intervals about a known zero in which the solution $W(x)$ does not vanish again.

THEOREM 5.1. *If $W(x)$ is a solution of (2.1), $W(0) = 0$, and if $\operatorname{Max} q_1(x) = M > 0$ in $0 \leqq x \leqq a$, then $W(x) \neq 0$ in $0 < x \leqq \operatorname{Min}(a, \pi M^{-\frac{1}{2}})$ provided that $Q(x) \not\equiv M$, and $W(x) \neq 0$ in $0 < x < \pi M^{-\frac{1}{2}}$ and $Q(x) \equiv M$ provided that $\pi M^{-\frac{1}{2}} \leqq a$ and $W(x)$ has the zero $\pi M^{-\frac{1}{2}}$.*

Proof. Let $Y(x)$, $y(x)$ and $u(x)$ be the solutions, respectively, of the following differential systems:

---

[2] E. L. Ince, *Differential Equations* (Dover, 1944), p. 225.

$$Y'' + (q_1(x) - \Phi^2(x))Y = 0, \qquad Y(0) = 0; \tag{5.1}$$

$$y'' + q_1(x)y = 0, \qquad y(0) = 0; \tag{5.2}$$

$$u'' + Mu = 0, \qquad u(0) = 0, \tag{5.3}$$

where $Y(x)$ is the associated solution of $W(x)$, and $y(x) \not\equiv 0$, $u(x) \not\equiv 0$. Denote by b, c, and d, respectively, the smallest positive zeros of $Y(x)$, $y(x)$ and $u(x)$. If $\Phi^2(x) \not\equiv 0$ and $q_1(x) \not\equiv M$, an appeal to Sturm's comparison theorem gives $0 < d < c < b$. Hence $c = b$ implies $\Phi^2(x) \equiv 0$ and $d = c$ implies $q_1(x) \equiv M$ in $0 \leqq x \leqq a$. Since $d = \pi M^{-\frac{1}{2}}$, if $W(x)$ has a zero $\pi M^{-\frac{1}{2}}$, then $d = c = b$ and hence $\Phi^2(x) \equiv 0$, $q_1(x) \equiv M$. Consequently from (2.8), $q_2(x) \equiv 0$ and $Q(x) \equiv q_1(x) \equiv M$. The rest of the theorem is now obvious.

THEOREM 5.2. *If $W(x)$ is a solution of (2.1), $W(0) = 0$, and if $q_1(x) \leqq 0$ in $0 \leqq x \leqq a$ and $q_1(x)$ changes sign at $x = a$, finally if $\operatorname{Max} q_1(x) = M > 0$ in $a \leqq x \leqq b$, then $W(x) \neq 0$ in $0 < x \leqq \operatorname{Min}(b, a + 2^{-1}\pi M^{-\frac{1}{2}})$.*

Proof. $Y(x)$ is a solution of the differential system

$$Y'' + (q_1(x) - \Phi^2(x))Y = 0, \qquad Y(0) = 0. \tag{5.4}$$

Define $y(x) = Cx$, $C > 0$, in $0 \leqq x \leqq a$, and $y(x) = A\cos(M^{\frac{1}{2}}x + B)$ in $a < x \leqq b$, where $Ca = A\cos(M^{\frac{1}{2}}a + B)$, $C = -A\sqrt{M}\sin(M^{\frac{1}{2}}a + B)$. It is easy to see that $y(x)$ satisfies the following differential system, except at $x = a$,

$$y'' + f(x)y = 0, \qquad y(0) = 0, \tag{5.5}$$

where $f(x) = 0$ in $0 \leqq x \leqq a$, $f(x) = M$ in $a < x \leqq b$. Following a line of reasoning similar to that in the proof of Sturm's comparison theorem, we find that the smallest positive zero of $Y(x)$ is greater than the smallest positive zero of $y(x)$. If the latter is denoted by c, then $(M^{\frac{1}{2}}c + B) - (M^{\frac{1}{2}}a + B) > \pi/2$, hence $c > a + 2^{-1}\pi M^{-\frac{1}{2}}$.

Similar zero-free intervals to the left of the origin can be determined by using Theorems 5.1 and 5.2 after obvious modifications.

THEOREM 5.3. *If $W(x)$ is a solution of (2.1), $W(0)W'(0) = 0$, and if $q_2(x)$ keeps one sign in $a \leqq x \leqq 0$, and in $0 \leqq x \leqq b$, then $W(x)W'(x) \neq 0$ in $a \leqq x < 0$, $0 < x \leqq b$.*

Proof. This theorem follows immediately from Theorem 4.2.

6. Along analytic curves, linear segments. If p is a point on an

analytic curve [3] C, $x = f(t)$, $y = g(t)$, $a \leqq t \leqq b$, then p can be represented by $z = x + iy = f(t) + ig(t)$. Along C the differential equation (1.1) has the form

$$d^2W/dt^2 + A(t)dW/dt + B(t)W = 0. \tag{6.1}$$

Let $H(t) = W(t)\exp(2^{-1}\int A(t)dt)$. Then equation (6.1) reduces to

$$d^2H/dt^2 + J(t)H = 0. \tag{6.2}$$

Hence our results obtained on the real axis can be extended immediately to the zeros along any analytic curve.

Along the ray $\theta = \theta_0$, any point p can be represented by $z = re^{i\theta_0}$, $0 \leqq \theta_0 < 2\pi$. Along this ray, (1.1) becomes

$$d^2W/dr^2 + Q(re^{i\theta_0})e^{i2\theta_0}W = 0. \tag{6.3}$$

The real and imaginary parts of $Q(r\exp(i\theta_0))\exp(i2\theta_0)$ are, respectively, $A = q_1(r, \theta_0)\cos 2\theta_0 - q_2(r, \theta_0)\sin 2\theta_0$, $B = q_1(r, \theta_0)\sin 2\theta_0 + q_2(r, \theta_0)\cos 2\theta_0$, where $Q(r\exp(i\theta_0)) = q_1(r, \theta_0) + iq_2(r, \theta_0)$, $q_1(r, \theta_0)$ and $q_2(r, \theta_0)$ are real.

Let $W(z)$ vanish at the origin. The theorems of section 5 provide us methods to determine linear segments on the ray $\theta = \theta_0$ in which $W(z)$ does not vanish again. Let r_0 be the positive root of the equation

$$\pi\{\mathrm{Max}\, A(r, \theta_0)\}^{-\frac{1}{2}} = r_0, \text{ where } 0 \leqq r \leqq r_0. \tag{6.4}$$

Then by Theorem 5.1, $W(z) \neq 0$ on the ray $\theta = \theta_0$ for $0 < r < r_0$ or $0 < r \leqq r_0$ according as $Q(r\exp(i\theta_0))\exp(i2\theta_0)$ is a positive constant or not. Denote this linear segment by $L_1(\theta_0)$. If $A(r, \theta_0) \leqq 0$ for all $r \geqq 0$, let $L_1(\theta_0)$ be the ray $0 < r < \infty$, $\theta = \theta_0$. If $A(r, \theta_0) \leqq 0$ for $0 \leqq r \leqq r_1$ and $A(r, \theta_0)$ changes sign at $r = r_1$, and if r_2 is the root of the equation

$$2^{-1}\pi\{\mathrm{Max}\, A(r, \theta_0)\}^{-\frac{1}{2}} = r_2 - r_1, \text{ where } r_1 \leqq r \leqq r_2, \tag{6.5}$$

then by Theorem 5.2, $W(z) \neq 0$ on the ray $\theta = \theta_0$ for $0 < r \leqq r_2$. Denote this linear segment by $L_2(\theta_0)$. Again if $A(r, \theta_0) \leqq 0$ for all $r \geqq 0$, let $L_2(\theta_0)$ be the ray $0 < r < \infty$, $\theta = \theta_0$. If $B(r, \theta_0)$ keeps one sign for $0 < r < r_1$ on the ray $\theta = \theta_0$ and changes sign at $r = r_1$, then by Theorem 5.3 $W(z)W'(z)$ does not vanish along the segment $0 < r \leqq r_1$, $\theta = \theta_0$. Denote this linear segment by $L_3(\theta_0)$. If $B(r, \theta_0) \equiv 0$, we agree that $L_3(\theta_0)$ contains no points.

7. Zero-free regions. The results of section 6 enable us to construct regions of the z-plane in which $W(z)$ does not vanish.

---

[3] See W. F. Osgoods' Lehrbuch der Funktionentheorie I, p. 702.

As θ_0 varies from 0 to 2π, each linear segment $L_i(\theta_0)$, i $= 1, 2, 3$, generates a region which we denote by R_i. In view of the discussion in section 6 and the construction of R_i, the following theorem is now obvious.

THEOREM 7.1. *If $W(z)$ is a solution of (1.1), $W(0) = 0$, then $W(z) \neq 0$ in $R_1 + R_2 + R_3$, $W(z)W'(z) \neq 0$ in R_3.*

THEOREM 7.2. *If $W(z)$ is a solution of (1.1), $W(0) = 0$, and if $\operatorname{Max} |Q(z)| = M > 0$ in $|z| \leqq a$, then $W(z) \neq 0$ in $0 < |z| \leqq \operatorname{Min}(a, \pi M^{-\frac{1}{2}})$ if $Q(z) \not\equiv$ const., $W(z) \neq 0$ in $0 < |z| < \pi M^{-\frac{1}{2}}$ and it has exactly two zeros on the circle $|z| = \pi M^{-\frac{1}{2}}$ and $Q(z) = C$, $|C| = M$, if it has a zero on this circle and $\pi M^{-\frac{1}{2}} \leqq a$.*

Proof. The differential equation (1.1) along a fixed ray reduces to equation (6.3). Since

$$M = \operatorname*{Max}_{|z| \leqq a} |Q(z)| = \operatorname*{Max}_{|z| \leqq a} |Q(re^{i\theta_0})e^{i2\theta_0}| \geqq \operatorname*{Max}_{0 \leqq r \leqq a} A(r, \theta_0),$$

where $A(r, \theta_0)$ is the real part of $Q(r\exp(i\theta_0))\exp(i2\theta_0)$, hence by Theorem 5.1, we have $W(z) \neq 0$ on the ray for $0 < r \leqq \operatorname{Min}(a, \pi M^{-\frac{1}{2}})$ provided $Q(z) \not\equiv$ constant and $\operatorname{Max} A(r, \theta_0)$ for $0 \leqq r \leqq a$ is positive. If $A(r, \theta_0) \leqq 0$ in $0 \leqq r \leqq a$, the assertion above holds in view of Theorem 5.2. Letting θ_0 vary from 0 to 2π, we have the first part of the theorem.

Let $\pi M^{-\frac{1}{2}} \leqq a$ and let $W(z)$ have a zero on the circle $|z| = \pi M^{-\frac{1}{2}}$, say $\pi M^{-\frac{1}{2}}\exp(i\theta_1)$, $\theta_1 =$ constant. Applying Theorem 5.1 to equation (6.3) gives $Q(r\exp(i\theta_1))\exp(i2\theta_1) \equiv M$. Hence $Q(z)$ is a constant on the ray $\theta = \theta_1$. But $Q(z)$ is analytic, so $Q(z) = M\exp(-2i\theta_1) = C$, $|C| = M$. The second part of the theorem is now obvious.

Along a straight line parallel to the imaginary axis, $z = x_0 + iy$, equation (1.1) reduces to

$$d^2W/dy^2 - Q(x_0 + iy)W = 0. \tag{6.6}$$

If $Q(x + iy) = q_1(x, y) + iq_2(x, y)$, $W(x + iy) = u(x, y) + iv(x, y)$ and $Y(x, y) = \pm (u^2(x, y) + v^2(x, y))^{\frac{1}{2}}$, where $q_1(x, y)$, $q_2(x, y)$, $u(x, y)$ and $v(x, y)$ are real, then $Y(x_0, y)$ satisfies

$$\partial^2 Y/\partial y^2 - (q_1(x_0, y) + \Phi^2(x_0, y))Y = 0. \tag{6.7}$$

The following lemma is used to prove the next theorem.

LEMMA. *If $W(z)$ is a solution of (1.1) and is real on the real axis, and if $q_1(x_0, y) \geqq 0$ for $0 \leqq y \leqq a$ and $q_1(x_0, y)$ changes sign at $y = a$,*

finally if $\text{Max}\{-q_1(x_0, y)\} = M$ *for* $a \leqq y \leqq b$, *then* $W(x_0 + iy) \neq 0$ *for* $0 < y \leqq \text{Min}(b, a + 2^{-1}\pi M^{-\frac{1}{2}})$.

Proof. Using the Cauchy-Riemann conditions and condition (2), it is easy to show $Y(x_0, 0)\partial Y(x_0, 0)/\partial y = 0$. Hence either $Y(x_0, 0) = 0$ or $\partial Y(x_0, 0)/\partial y = 0$, but not both.

Case I. $Y(x_0, 0) = 0$, $\partial Y(x_0, 0)/\partial y \neq 0$. The lemma follows from a proof similar to that of Theorem 5. 2.

Case II. $Y(x_0, 0) \neq 0$, $\partial Y(x_0, 0)/\partial y = 0$. Define $S(y)$ *to be* $Y(x_0, 0)$ in $0 \leqq y \leqq a$, and $Y(x_0, 0)\cos(M^{\frac{1}{2}}y + D)$ in $a < y \leqq b$, where $\cos(M^{\frac{1}{2}}a + D) = 1$, $\sin(M^{\frac{1}{2}}a + D) = 0$. Then $S(y)$ satisfies the following equation, except at $y = a$,

$$d^2S/dy^2 + f(y)S = 0, \tag{6.8}$$

where $f(y) = 0$ for $0 \leqq y \leqq a$, $f(y) = \text{M}$ for $a < y \leqq b$. On the other hand $Y(x_0, y)$ is a solution of equation (6. 7). Along a line of reasoning similar to that of the proof of Sturm's comparison theorem, it is seen that the smallest positive zero of $S(y)$ is less than the smallest positive zero of $Y(x_0, y)$. The smallest positive zero of $S(y)$ is $a + 2^{-1}\pi M^{-\frac{1}{2}}$. Hence the lemma is established.

After obvious modifications, the lemma applies also to the other half of the line $z = x_0 + iy$.

Let $q_1(x_0, y) \geqq 0$ for $0 \leqq y \leqq a$ and suppose it changes sign at $y = a$. Let b be a root of $2^{-1}\pi\{\text{Max} - q_1(x_0, y)\}^{-\frac{1}{2}} = b - a$, where $a \leqq y \leqq b$, and denote by $P(x_0)$ the point $z = x_0 + ib$. Using a similar method, determine a point $Q(x_0)$ on the line $z = x_0 + iy$ below the real axis. The lemma shows that $W(z)$ has no complex zero along the linear segment $P(x_0)Q(x_0)$. Letting x_0 vary from α to β, the segment $P(x_0)Q(x_0)$ generates a region which is denoted by $R_{\alpha\beta}$.

THEOREM 7. 3. *If* $W(z)$ *is a solution of* (1. 1) *which is real on the real axis, then* $W(z)$ *has no complex zero in* $R_{\alpha\beta}$, *where* $R_{\alpha\beta}$ *is any of the regions constructed above.*

This theorem, which follows from the construction of $R_{\alpha\beta}$, holds also if $W(z)$ is imaginary, instead of real, on the real axis, because $iW(z)$ is then real.

If $W(z)$ is real or imaginary on the imaginary axis, a similar zero-free region may be constructed.

THE UNIVERSITY OF MISSOURI.

ON THE EQUATION $a^x - b^y = 1$.*

By Wm. J. LeVeque.

1. Introduction. Equations similar to that of the title were apparently first considered by S. S. Pillai [1][1] in 1931. He showed then that for any fixed set of integers a, b, c, the equation $a^x - b^y = c$ has only finitely many solutions, and that for fixed a, b the number of solutions of the inequality $0 < a^x - b^y \leqq n$ is asymptotic to

$$\frac{\log^2 n}{2 \log a \log b}$$

as $n \to \infty$. A. Herschefeld [2] showed that the equation $2^x - 3^y = c$ has at most one solution for large c, and Pillai [3] extended this to $a^x - b^y = c$ for $c > c_0 = c_0(a, b)$. We show here that the equation

$$a^x - b^y = 1 \tag{1}$$

has at most one solution unless $a = 3$, $b = 2$, when there are just two. The complete statement is contained in Theorem 6.

Since there is evidently no solution of (1) if $(a, b) > 1$, we assume throughout that a and b are relatively prime and larger than unity. Then exactly one of a, b is even; the case with a even is discussed in § 2, the case with a odd in § 3. Lemma 2 of § 3 was proved by A. Bindoni [4], but the proof given here is somewhat shorter and is included for completeness.

As application of the principal result, it is shown in § 4 that the relation $\sum_{k=1}^{n} k^3 = (\sum_{k=1}^{n} k)^2$ is the only one of its kind, i. e., that it is the only instance of the equation

$$\sum_{k=1}^{n} k^s = (\sum_{k=1}^{n} k^t)^r$$

with $r > 1$.

All letters stand for rational integers. If α belongs to $f \pmod{\beta}$ (i. e., if $z = f$ is the smallest positive solution of the congruence $\alpha^z \equiv 1 \pmod{\beta}$), we write $\alpha \to f(\beta)$. The symbol $[t_1, \cdots, t_r]$ means the l. c. m. of $t_1, \cdots, t_r$. If p is prime and $p^\alpha \mid \beta$ while $p^{\alpha+1} \nmid \beta$, we write $p^\alpha \parallel \beta$.

* Received May 8, 1950.

[1] Numbers in brackets refer to the bibliography at the end of this article.

2. The case a even, b odd.

THEOREM 1. *The equation* (1) *has at most one solution if a is even and b is odd. If $2^\alpha \| a$ and $2^\beta \| b+1$, the only possible solution is with $x = \beta/\alpha$, if $\alpha > 1$, and with $x = 1$ or β if $\alpha = 1$.*

First assume that (1) holds and that at least one of α, x is larger than unity; i. e., that $\alpha x \geqq 2$. Then

$$a^x = (b-1+1)^y + 1 = (b-1)^y + y(b-1)^{y-1} + \cdots + y(b-1) + 2.$$

Put $a = 2^\alpha \cdot a_1$:

$$2^{\alpha x - 1} a_1{}^x = 2^{y-1}\left(\frac{b-1}{2}\right)^y + y2^{y-2}\left(\frac{b-1}{2}\right)^{y-1} + \cdots + y\frac{b-1}{2} + 1;$$

it follows that y is odd. Putting $b = 2^\beta \cdot b_1 - 1$, we have $a^x = (2^\beta b_1 - 1)^y + 1$, so that $2^\beta \| a^x$, $\beta = \alpha x$, $x = \beta/\alpha$.

The proof will be complete when we show that if $\alpha = 1$ and (1) has a solution with $x = 1$, it has no solution with $x > 1$. Assume that there is a y_0 such that $a - b^{y_0} = 1$, and that (1) holds for some $x > 1$. Then $y > y_0$ and we have

$$(b^{y_0}+1)^x - b^y = 1, \qquad b^{y_0 x} + xb^{y_0(x-1)} + \cdots + xb^{y_0} - b^y = 0,$$

whence $b \mid x$. Put $x = bx'$. By the earlier part of the proof,, the only possible solution of $(a^b)^{x'} - b^y = 1$ is with $x' = \beta/\alpha'$, where $\alpha' = \alpha b = b$. But $\beta < \alpha'$ for $b > 1$, so that $\alpha' \uparrow \beta$.

3. The case a odd, b even.

THEOREM 2. *If $a \to 1(b)$ and $b \to 2(a)$, then $a = b+1$.*

We have $a = jb+1$, $b^2 = ha+1$, so that $b^2 - jhb - (h+1) = 0$. The discriminant of this quadratic must be a square, so that $j^2h^2 + 4h + 4 = r^2$. Since $r > jh$ we can put $r = jh + \alpha$ with $\alpha > 0$. This gives $4h + 4 = 2\alpha jh + \alpha^2$, so that α is even, and $\alpha \leqq 2$ or h is negative. Since h is positive, $\alpha = 2$, $j = 1$, $a = b+1$.

Now assume that there is a solution of (1), where now a is odd and b is even. Then clearly b belongs to an even exponent (mod a), and we define s_0, t_0 by the conditions $a \to s_0(b)$, $b \to 2t_0(a)$, and define s_k, t_k recursively by the conditions

$$a^{s_0 \ldots s_{k-1}} \to s_k(b^{t_0 \ldots t_{k-1}}), \qquad b^{t_0 \ldots t_{k-1}} \to 2t_k(a^{s_0 \ldots s_{k-1}}).$$

THEOREM 3. *If* (1) *is solvable, then there is a g for which $x_0 = s_0 \cdots s_g$,*

$y_0 = t_0 \cdot \cdots t_g$ *is a solution, and for every solution* (x, y) of (1), $x_0 \mid x$ *and* $y_0 \mid y$.

Let (x, y) be a solution of (1). Then clearly $s_0 \mid x$, $t_0 \mid y$; put $x_1 = x/s_0$, $y_1 = y/t_0$. If $s_0 = t_0 = 1$, then by Theorem 2, $a - b = 1$, and the theorem holds with $g = 0$. If $s_0 t_0 > 1$, then $x_1 y_1 < xy$ and we have $(a^{s_0})^{x_1} - (b^{t_0})^{y_1} = 1$. As before, $s_1 \mid x_1$, $t_1 \mid y_1$. If $s_1 = t_1 = 1$ the theorem holds with $g = 0$; otherwise, putting $x_2 = x_1/s_1$, $y_2 = y_1/t_1$, we have $(a^{s_0 s_1})^{x_2} - (b^{t_0 t_1})^{y_2} = 1$, where $x_2 y_2 < x_1 y_1$. Continuing in this fashion we must encounter a k for which $s_k = t_k = 1$, for otherwise we would have an infinite decreasing sequence of positive integers $x_i y_i$. Then $g = k - 1$.

Two problems now confront us: to investigate the g of Theorem 3 and to show that unless a is 3 and b is 2, (x_0, y_0) is the only possible solution of (I). We need some preliminary lemmas.

LEMMA 1. *If* p *is an odd prime and* $p \nmid ab$, *each of the congruences* $a \equiv b \pmod{p^k}$ *and* $a^{p^s} \equiv b^{p^s} \pmod{p^{k+s}}$ *implies the other.*

That the first congruence implies the second is easily proved by induction on s. The reverse implication is proved for $s = 1$ by using Fermat's theorem and induction on k, and this result is used to prove the general statement by induction on s.

LEMMA 2. *If* $\alpha \to t(p)$, p *an odd prime, and if* $p^z \parallel \alpha^t - 1$, *then* $\alpha \to t p^{m(n-z)}(p^n)$, *where* $m(x)$ *denotes* x *for* $x > 0$ *and* 0 *for* $x \leqq 0$.

Assume the hypotheses of the theorem are satisfied. If $n \leqq z$, then $p^n \mid \alpha^t - 1$. If $p^n \mid \alpha^{t'} - 1$, then $p \mid \alpha^{t'} - 1$, so that $t' \geqq t$. This proves the theorem for $n \leqq z$.

If $n > z$, Lemma 1 shows that $\alpha^{t p^{n-z}} \equiv 1 \pmod{p^n}$. We must show that $\alpha^d \not\equiv 1 \pmod{p^n}$ if d is a proper divisor of $t p^{n-z}$. Let $d = t_1 p^r$ with $r \leqq n - z$, $t_1 \mid t$, and assume that $\alpha^{t_1 p^r} \equiv 1 \pmod{p^n}$. By Lemma 1, $\alpha^{t_1} \equiv 1 \pmod{p^{n-r}}$. Since $n - r \geqq z$, this contradicts the definitions of t and z, unless $r = n - z$, $t = t_1$.

LEMMA 3. *If* $\alpha_i \to t_i(\beta_i)$ *for* $i = 1, \cdots, r$ *and if the* β_i *are pairwise coprime, then*

$$\prod_{i=1}^{r} \alpha_i \to [t_1, \cdots, t_r] \pmod{\prod_{i=1}^{r} \beta_i}.$$

This is obvious.

We have shown that any solution (x, y) of (1) is such that $x_0 \mid x, y_0 \mid y$. Putting $a^{x_0} = a_1$, $b^{y_0} = b_1$, $x/x_0 = x'$, $y/y_0 = y'$, this means that if (x, y) is a solution of (1), then (x', y') is a solution of $(b_1 + 1)^{x'} - b_1^{y'} = 1$. Theorem 4 shows that (x, y_0) is the only solution of (1) unless $a = 3$, $b = 2$.

THEOREM 4. *The only solutions of* $3^x - 2^y = 1$ *are* $(1, 1)$ *and* $(2, 3)$. *The only solution of* $(\alpha + 1)^x - \alpha^y = 1$ *with* α *even,* $\alpha > 2$ *is* $(1, 1)$.

Let p be an odd prime dividing $\alpha + 1$, and put $\alpha + 1 = mp^\gamma$, $\gamma \geqq 1$, $p \nmid m$. If $(\alpha + 1)^x - \alpha^y = 1$, then $(mp^\gamma - 1)^y \equiv -1 \pmod{p^{\gamma x}}$. Since $\alpha \to 2 (p)$ and $p^\gamma \parallel \alpha^2 - 1$, Lemma 2 shows that $\alpha \to 2p^{\gamma(x-1)} (p^{\gamma x})$. Hence $y \equiv p^{\gamma(x-1)} \pmod{2p^{\gamma(x-1)}}$, and so $y = Np^{\gamma(x-1)}$. This gives

$$(\alpha + 1)^x - \alpha^{Np^{\gamma(x-1)}} = 1.$$

But $2^{3^{x-1}} > 3^x - 1$ for $x > 2$, and for any $\alpha > 2$, $\alpha^{3^{x-1}} > (\alpha + 1)^x - 1$ for $x > 1$.

We must now investigate the g of Theorem 3.

THEOREM 5. $g = 0$ *or* 1. *In either case,* $y_0 = t_0$, $x_0 = s_0 s_1$.

Let $a = \Pi\, p_\rho^{\alpha_\rho}$; here and hereafter ρ ranges from 1 to r in all products, and $[f_\rho]$ will mean the l. c. m. of $f_1, \cdots, f_r$. We are assuming that (x_0, y_0) actually is a solution of (1), so that $b^{y_0} \equiv -1 \pmod{p_\rho}$ and the exponent to which b belongs (mod p_ρ) must be even, say $2\tau_\rho$. Let $p_\rho^{z_\rho} \parallel b^{2\tau_\rho} - 1$; then $p_\rho^{z_\rho} \parallel b^{\tau_\rho} + 1$. By Lemma 2,

$$b \to 2\tau_\rho p_\rho^{m(\alpha_\rho - z_\rho)} \pmod{p_\rho^{\alpha_\rho}},$$

so that, by Lemma 3, $t_0 = [\tau_\rho p_\rho^{m(\alpha_\rho - z_\rho)}]$. Now let $b \to 2\sigma_\rho\, (p_\rho^{\alpha_\rho x_0})$, for $\rho = 1, \cdots, r$. From the way it was defined, it is clear that y_0 is the smallest exponent such that $b^{y_0} \equiv -1 \pmod{a^{x_0}}$, so that $y_0 = [\sigma_\rho]$. From Lemma 2,

$$\sigma_\rho = \tau_\rho p_\rho^{m(\alpha_\rho x_0 - z_\rho)}, \qquad \rho = 1, \cdots, r$$

so that

$$y_0 = [\tau_\rho p_\rho^{m(\alpha_\rho x_0 - z_\rho)}] = t_0 \Pi\, p_\rho^{m(\alpha_\rho x_0 - z_\rho) - m(\alpha_\rho - z_\rho)},$$

and

$$a^{x_0} - b^{t_0 \Pi p_\rho^{m(\alpha_\rho x_0 - z_\rho) - m(\alpha_\rho - z_\rho)}} = 1. \tag{2}$$

Since $p_\rho^{z_\rho} \mid b^{\tau_\rho} + 1$, both $p_\rho^{\alpha_\rho}$ and $p_\rho^{z_\rho}$ divide $b^{t_0} + 1$. Let $p_\rho^{\zeta_\rho} \parallel b^{t_0} + 1$; then

$$\zeta_\rho \geqq \text{Max}(\alpha_\rho, z_\rho) \tag{3}$$

and for some κ noto divisible by any p_ρ,

$$b^{t_0} = \kappa \prod p_\rho^{\zeta_\rho} - 1. \tag{4}$$

Thus (2) gives

$$a^{x_0} = (\kappa \prod p_\rho^{\zeta_\rho} - 1)^{\prod p_\rho^{m(\alpha_\rho x_0 - z_\rho) - m(\alpha_\rho - z_\rho)}} + 1. \tag{5}$$

By Lemma 1, for each ρ,

$$(\kappa \prod p_\rho^{\zeta_\rho} - 1)^{p_\rho^{m(\alpha_\rho x_0 - z_\rho) - m(\alpha_\rho - z_\rho)}} \equiv -1 \pmod{p_\rho^{\zeta_\rho + m(\alpha_\rho x_0 - z_\rho) - m(\alpha_\rho - z_\rho)}},$$

so that (5) gives

$$p_\rho^{\alpha_\rho x_0} \equiv 0 \pmod{p_\rho^{\zeta_\rho + m(\alpha_\rho x_0 - z_\rho) - m(\alpha_\rho - z_\rho)}},$$

whence $\alpha_\rho x_0 \geqq \zeta_\rho + m(\alpha_\rho x_0 - z_\rho) - m(\alpha_\rho - z_\rho)$. But by (3) and the definition of $m(u)$,

$$\zeta_\rho + m(\alpha_\rho x_0 - z_\rho) - m(\alpha_\rho - z_\rho) \geqq \begin{cases} \alpha_\rho x_0 \text{ always} \\ \alpha_\rho x_0 + 1 \text{ if } z_\rho > \alpha_\rho x_0. \end{cases}$$

Hence

$$\zeta_\rho = \text{Max}(\alpha_\rho, z_\rho) \text{ and } z_\rho \leqq \alpha_\rho x_0. \tag{6}$$

For brevity, put $\pi = \prod p_\rho^{m(\alpha_\rho x_0 - z_\rho) - m(\alpha_\rho - z_\rho)}$ and put

$$\pi_{\rho_1} = \prod_{\substack{\rho=1 \\ \rho \neq \rho_1}}^{r} p_\rho^{m(\alpha_\rho x_0 - z_\rho) - m(\alpha_\rho - z_\rho)}.$$

We will show that unless $a = 3$, $b = 2$ the only solution of (5) is with $x_0 = s_0 s_1$, by showing that otherwise the right side of (5) is larger than the left when $\kappa = 1$, and *a fortiori* when $\kappa > 1$. We have

$$(\prod p_\rho^{\zeta_\rho} - 1)^\pi + 1 > \prod (p_\rho^{\zeta_\rho} - 1)^\pi$$

and wish to investigate the circumstances under which

$$p_\rho^{\alpha_\rho x_0} \geqq (p_\rho^{\zeta_\rho} - 1)^\pi. \tag{7}$$

Case I. For some ρ, $\alpha_\rho = z_\rho$. By (6), $\zeta_\rho = \alpha_\rho \geqq 1$. If we choose $p_\rho = 3$, $\alpha_\rho = 1$, $x_0 = 2$, $\pi_\rho = 1$, then

$$p_\rho^{\alpha_\rho x_0} = (p_\rho^{\alpha_\rho} - 1)^{p_\rho^{\alpha_\rho x_0 - \alpha_\rho} \cdot \pi_\rho} + 1.$$

Clearly any increase in p_ρ, α_ρ, x_0 or π_ρ would make

$$p_\rho^{\alpha_\rho x_0} < (p_\rho^{\alpha_\rho} - 1)^{p_\rho^{\alpha_\rho (x_0 - 1)} \pi_\rho} + 1,$$

and of these four quantities only x_0 can be decreased, so that the only solution of (7) with $x_0 > 1$ is that given above, which leads to the exception $3^2 - 2^3 = 1$. But if $x_0 = 1$ then $s_0 \cdot \cdot \cdot s_g = 1$, so that $s_0 = s_1 = 1$, from which it follows that $t_1 = 1$, so that $g = 0$.

Case II. For all ρ, $z_\rho > \alpha_\rho$. Then $\zeta_\rho = z_\rho \leqq \alpha_\rho x_0$ by (6). The only circumstances under which the inequality

$$p_\rho^{\alpha_\rho x_0} < (p_\rho^{z_\rho} - 1)^{p_\rho^{\alpha_\rho x_0 - z_\rho} \cdot \pi_\rho}$$

is false are that $p_\rho = 3$ or $\alpha_\rho x_0 - z_\rho = 0$, $\pi_\rho = 1$. If the only prime factor of a is 3, it is easy to see that again $\alpha_\rho x_0 - z_\rho = 0$. From (2), $a^{s_0 \cdots s_g} - b^{t_0} = 1$. Since $t_1 = 1$, $s_2 = 1$, so that $g = 0$ or 1. This completes the proof of Theorem 5.

Combining Theorems 1-5, we summarize:

THEOREM 6. *The equation* (1) *has just the two solutions* $(1, 1)$ *and* $(2, 3)$ *if* $a = 3$, $b = 2$. *In all other cases it has at most one solution. If* a *is even and* b *is odd, the only possible solution is with* $x = \beta/\alpha$ *if* $2^\alpha \parallel a$, $2^\beta \parallel b + 1$ *and* $\alpha > 1$, *and with* $x = 1$ *or* β *if* $\alpha = 1$. *The only possible solution if* a *is odd and* b *is even is* $(s_0 s_1, t_0)$, *where* $a \to s_0(b)$, $b \to 2t_0(a)$, $a^{s_0} \to s_1(b^{t_0})$.

4. An application. As an immediate consequence of Theorem 6, we have

THEOREM 7. *The equation*

$$b^s + 1 = (b^t + 1)^r \tag{8}$$

has no solution with $r > 1$ *if* $b > 2$; *it has the sole solution* $r = 2$, $s = 3$, $t = 1$ *if* $b = 2$.

For putting $a = b^t + 1$, (8) gives $a^r = b^s + 1$. But the equation $a^x - b^y = 1$ has the solution $(1, t)$, and therefore has no other solution unless $b = 2$, $a = 3$.

THEOREM 8. *The only solution of the identity* $\sum_{k=1}^{n} k^s = (\sum_{k=1}^{n} k^t)^r$ *with* $r > 1$ *is* $r = 2$, $s = 3$, $t = 1$.

For by taking $n = 2$ the identity reduces to equation (8) with $b = 2$.

UNIVERSITY OF MICHIGAN.

BIBLIOGRAPHY.

[1] S. S. Pillai, "On the inequality $0 < a^x - b^y \leqq n$," *Journal of the Indian Mathematical Society*, vol. 19 (1931), pp. 1-11.

[2] A. Herschefeld, "The equation $2^x - 3^y = d$," *Bulletin of the American Mathematical Society*, vol. 42 (1936), pp. 231-234.

[3] S. S. Pillai, "On $a^x - b^y = c$," *Journal of the Indian Mathematical Society*, New Series, vol. 2 (1936), pp. 119-122.

[4] A. Bindoni, "A proposito di un articolo del Prof. Loria sui caratteri de divisibilita," *Giorn. scientifico-didattico per l'incremente degli studi Mathmatici nelle Scuole Medie*, vol. 4 (1905), pp. 87-92.

SUM FUNCTIONS IN ELEMENTARY p-ADIC ANALYSIS.*

By GORDON OVERHOLTZER.

Introduction. I. Schur in [5] introduced as the derivate of the sequence $\{a_n\}$ with respect to the number p the sequence $\{\Delta a_n\}$, the individual term of which is given by

$$\Delta a_n = (a_{n+1} - a_n)/p^{n+1}, \tag{0, 1}$$

in order to generalize Fermat's theorem of elementary number theory. Fermat's theorem states that if p is a rational prime, a a rational integer, the rational numbers Δa^{p^n} are integers. Schur proved that the terms of the first $p-1$ Schur derivates of $\{a^{p^n}\}$ are integral if $(a, p) = 1$, where (a, p) is the greatest common divisor of a and p.

The author in [4] used the Schur derivates to generalize the number theoretical result that the sum of the k-th powers (k an integer) of the integers less than and prime to p^n, p a rational prime, n a positive integer, is divisible by p^{n-1} or p^n according as $p-1$ does or does not divide k. Thus the sequence

$$\{[(i, p) = 1, 1 \leqq i \leqq p^n] \sum_i i^k/p^n\}, \tag{0, 2}$$

has denominators at most p. (Summation conditions will frequently be written before the summation sign. If the summation conditions are not given, they remain those previously indicated for that summation index.) It was shown that the Schur derivates of all orders of the sequence (0, 2) are p-adically bounded. This result followed from a general formula for the Schur derivates of the sequence

$$\{[(i, p) = 1, 1 \leqq i \leqq p^n] \sum_i g(i)/p^n\}, \tag{0, 3}$$

where g is a function defined on the p-adic integers satisfying certain analyticity conditions. It was noted that the function $g(x)$ might be defined by different power series in different residue classes.

In the present paper, sums of functions defined by power series where

* Received March 17, 1950. This paper forms a part of the author's doctoral dissertation written at Indiana University. The author wishes to take this opportunity to thank Professor M. A. Zorn for his aid and encouragement in the preparation of this paper.

the argument is restricted to a single residue class module p^e are studied. It is found that the Schur derivates of the sequence

$$(0,4) \qquad \{[i \equiv i_0 (\mathrm{mod}\, p^e), 0 \leqq i \leqq p^n - 1] \sum_i g(i)/p^n\},$$

formed by summation in a single residue class up to p^n and division by p^n, are given by formulae analogous to those for the derivates of $(0,3)$. The sum over several residue classes may be defined additively.

A part of the theory is extended to generalized representatives assigned to the non-negative integers by a recursion using the development of the integers in the number base p, given an initial set of representatives assigned to $0, \cdots, p-1$, and an element with the same valuation as p. The generalization of arguments over which sums are to be taken is desirable for discussion of the $\mathfrak{p}$-adic analogs of the p-adic theorems given here. The $\mathfrak{p}$-adic case is not discussed in this paper.

A general expression in terms of Bernoullian numbers is found for the p-adic limit of the sequence $(0,4)$, and evaluated for certain special functions $g(x)$. The limit for $g(x) = \exp x$, $i \equiv 0 \pmod p$ is evaluated by two methods. Equality of the two results yields the same equation as that given by a power series identity on substituting $x = p$, for which both members are defined p-adically.

The limit for $g(x) = \log x$, $i \equiv 1 \pmod p$ suggests the asymptotic Stirling series for $\log \Gamma(x)$, where $\Gamma(x)$ is the classical gamma-function. If Γ_p is defined for positive rational integers by $\Gamma_p(n) = \prod_{i=0}^{n-1} (1 + ip)$, the limit of the divided sums $(0,4)$ for $g(x) = \log x$ is given by $p^{-1} \lim \log \Gamma_p(p^n)/p^n$, where $n \to \infty$. The function $\Gamma_p(n)$ is shown to be uniformly continuous and hence capable of a uniformly continuous extension to the p-adic integers. The proof depends upon a generalization of Wilson's theorem.

The examples studied suggest consideration of the difference equation

$$(0,5) \qquad f(x+1) - f(x) = g(x),$$

with a given initial condition. If g is continuous, the difference equation $(0,5)$ has exactly one continuous solution in p-adic analysis, given by the continuous extension of $S(n,g) = \sum_{i=0}^{n-1} g(i)$, which is the function formed by summing $g(x)$ over the integers up to n. If $f(x)$ is given by a power series with somewhat restricted coefficients, the sum function $S(x,f)$ is analytic, that is, given by a power series. For generalized representatives for which the initial representative set is complete, the sum function of a continuous function is continuous.

Some limit results permit the interpretation that $\lim S(p^n, g)/p^n$, where $n \to \infty$, is the derivative of $S(x, g)$ at zero.

1. Representatives, sums, and Schur derivates. In this section, the manner of assignment of representatives to the non-negative integers is described. Notation for the sums of functions $g(x)$ with the argument restricted to the representatives of a single residue class modulo p^e is introduced. The Schur derivates of all orders of the sequence formed by summation in a single residue class up to p^n of the values of a power series, are given as infinite series in sums of values of the derivatives of the power series.

Throughout the paper p will denote a rational prime, and the analysis will be understood to be p-adic. In particular, $|a|$ will be understood to be the p-adic valuation of a. Let $r_0, \cdots, r_{p-1}$ be an arbitrary set of p-adic integers. Express the positive rational integer i in the number base p. If $i = \sum_{k=0}^{e} i_k p^k$, $0 \leqq i_k \leqq p-1$, and $i \geqq p$, define $r_i = \sum_{k=0}^{e} r_{i_k} q^k$, where $|q| = |p|$. A set of p-adic integers r_i assigned to the integers according to this definition is called a set of generalized representatives. If the set of initial representatives $r_0, \cdots, r_{p-1}$, forms a complete residue set, the representatives are said to be initially complete.

If, in particular, $r_i = i$, $i = 0, \cdots, p-1$, and $q = p$, then $r_i = i$ for all integers i. These representatives will be called standard.

(1, 1) THEOREM. *If a set of generalized representatives is initially complete, then* $r_{i'} \equiv r_{i''} \pmod{p^e}$ *if and only if* $i' \equiv i'' \pmod{p^e}$.

$$\text{Write } i' = \sum_{k=0}^{e'} i_k' p^k,\ i'' = \sum_{k=0}^{e''} i_k'' p^k,\ 0 \leqq i_k', i_k'' \leqq p-1.$$

Then if $r_{i'} \equiv r_{i''} \pmod{p^e}$,

$$\sum_{k=0}^{e-1} (r_{i_k'} - r_{i_k''}) q^k + \sum_{k=e}^{e'} r_{i_k'} q^k - \sum_{k=e}^{e''} r_{i_k''} q^k \equiv 0 \pmod{p^e}.$$

Since $|q| = |p|$, and the same congruence holds to the modulus p, $r_{i_0'} \equiv r_{i_0''} \pmod{p}$, and $r_{i_0'} = r_{i_0''}$ or $i_0' = i_0''$, since $r_0, \cdots, r_{p-1}$ form a complete residue set. Similarly $i_k' = i_k''$, $k = 1, \cdots, e-1$. Then $i' \equiv i'' \pmod{p^e}$.

If $i' \equiv i'' \pmod{p^e}$, $i_k' = i_k''$, $k = 0, \cdots, e-1$. Then

$$r_{i'} - r_{i''} = \sum_{k=e}^{e'} r_{i_k'} q^k - \sum_{k=e}^{e''} r_{i_k''} q^k \equiv 0 \pmod{p^e}.$$

If the assignment of representatives to the non-negative rational integers is

regarded as a mapping from the non-negative rational integers to the p-adic integers, (1, 1) states that r is an isometry, that is, $| i' - i'' | = | r_{i'} - r_{i''} |$.

Let $g(x)$ be a function defined for p-adic integers. Let us introduce an abbreviation for the sum of function values when the argument is restricted to representatives assigned to the integers in a single residue class modulo p^e and less than a given integer n. Define

$$(1,2) \qquad \sigma[i_0(p^e), n, g] = [i \equiv i_0 \pmod{p^e}, 0 \leqq i \leqq n-1] \sum_i g(r_i),$$

$$(1,3) \qquad \delta[i_0(p^e), n, g] = \sigma[i_0(p^e), n, g]/n.$$

Greek letters will indicate summation over generalized representatives; δ is a divided sum.

For summation over standard representatives, capital Roman letters will be used; thus we write

$$(1,4) \qquad S[i_0(p^e), n, g] = [i \equiv i_0 \pmod{p^e}, 0 \leqq i \leqq n-1] \sum_i g(i),$$

$$(1,5) \qquad D[i_0(p^e), n, g] = S[i_0(p^e), n, g]/n.$$

The definitions (1, 2) through (1, 5) are for a single residue class, the neighborhood of i_0 with radius $| p^e |$. For a finite set of neighborhoods, these restricted sums may be defined by summation of the appropriate formulae.

According to (0, 1), the individual terms of the Schur derivates of $\{\delta[i_0(p^e), p^n, g]\}$ with respect to q are given by

$$\begin{aligned}\Delta^m \delta[i_0(p^e), p^n, g] &= \Delta(\Delta^{m-1}\delta[i_0(p^e), p^n, g]) \\ &= (\Delta^{m-1}\delta[i_0(p^e), p^{n+1}, g] - \Delta^{m-1}\delta[i_0(p^e), p^n, g])/q^{n+1}.\end{aligned}$$

If $g(x)$ is given by a power series, and the divided sums $\delta[i_0(p^e), p^n, g^{(k)}]$ are bounded uniformly, the terms of the Schur derivates can be expressed as an infinite series in $\delta[i_0(p^e), p^n, g^{(k)}]$, where $g^{(k)}$ is the k-th derivative of g.

(1, 6) THEOREM. *If $g(x)$ is a function defined for $x \equiv r_{i_0} \pmod{p^e}$ and can be expressed in a power series*

$$g(x) = \sum_{j=0}^{\infty} g^{(j)}(r_{i_0})(x - r_{i_0})^j/j!,$$

convergent if $| x - r_{i_0} | \leqq | p^e |$, $\delta[i_0(p^e), p^n, g^{(k)}]$ is bounded uniformly with respect to k, and the Schur derivates are taken with respect to q, then there exist coefficients $L^{(m)}(\gamma)$ independent of n and g such that for $n \geqq e$,

$$(1,7) \qquad \Delta^m \delta[i_0(p^e), p^n, g] = \sum_{\gamma=m}^{\infty} L^{(m)}(\gamma) q^{n(\gamma-m)} \delta[i_0(p^e), p^n, g^{(\gamma)}],$$

$$(1,8) \qquad L'(\gamma) = (pq)^{-1} \sum_{i=0}^{p-1} (r_i)^\gamma / \gamma\,!,$$

$$(1,9) \qquad L^{(m+1)}(\gamma) = [\mu + \nu = \gamma, \mu \geqq m, \nu \geqq 1] \sum_{\mu,\nu} q^{\mu-m} L^{(m)}(\mu) L'(\nu) + L^{(m)}(\gamma)(q^{\gamma-m-1} - q^{-1}),$$

$$(1,10) \qquad |L^{(m)}(\gamma)| \leqq |p^{-[2m+\gamma/(p-1)]}|.$$

The proof is similar to the proof of Theorem (1, 8) of [4]. To construct the proof from the one given in [4] write $\gamma[i_0(p^e), p^n, g]$ for $T[n, g(x)]$, $\delta[i_0(p^e), p^n, g]$ for $S[n, g(x)]$, r_γ for γ as an argument of g, and q for p in the power series development and the formation of the Schur derivate. The divided sums are still formed with p, which plays here the role of the number of residue classes modulo p.

The exponent of p in $L^{(m)}(\gamma) q^{n(\gamma-m)} \delta[i_0(p^e), p^n, g^{(\gamma)}]$, $\gamma \geqq m$, is at least $n(\gamma - m) - 2m - \gamma/(p-1) + N$, if $|p^N|$ is the uniform bound of $\delta[i_0(p^e), p^n, g^{(\gamma)}]$. This exponent is an increasing function of γ. Therefore all terms of the m-th Schur derivate after the first contain a power of p with exponent at least $n - 2m - (m+1)/(p-1) + N$, the value of the exponent for $\gamma = m+1$. The exponent $n(\gamma - m) - 2m - \gamma/(p-1) + N$ is an increasing function of n. Hence $\{\Delta^m \delta[i_0(p^e), p^n, g]\}$ is bounded for all m, $\{\Delta^{m-1} \delta[i_0(p^e), p^n, g]\}$ is convergent for all m. Now

$$\lim_{n\to\infty} \Delta^m \delta[i_0(p^e), p^n, g] = \lim_{n\to\infty} \{L^{(m)}(m) \delta[i_0(p^e), p^n, g^{(m)}] + I p^{n-2m-(m+1)/(p-1)+N}\},$$

where $|I| \leqq 1$. Hence

$$(1,11) \qquad \lim_{n\to\infty} \Delta^m \delta[i_0(p^e), p^n, g] = L^{(m)}(m) \lim_{n\to\infty} \delta[i_0(p^e), p^n, g^{(m)}].$$

If the set of initial representatives $r_0, \cdots, r_{p-1}$ is the set consisting of zero and the $p-1$-st roots of unity, $L'(\gamma) = (pq)^{-1} \sum_{i=0}^{p-1} (r_i)^\gamma / \gamma\,! = 0$, if $p-1$ does not divide γ. From the recursion (1, 9) it follows that for all m, $L^{(m)}(\gamma) = 0$, if $p-1$ does not divide γ.

If the generalized representatives are initially complete, the requirement of uniform boundedness of $\delta[i_0(p^e), p^n, g^{(k)}]$ can be omitted from the hypothesis of Theorem (1, 6). The number-theoretical result

$$(1,12) \qquad \sum_{i=0}^{p^{n-1}} i^k \equiv 0 \pmod{p^{n-1}}, \; k \text{ a positive integer},$$

will be required. Since

$$\sum_{i=0}^{p^n-1} i^k = \sum_{j=0}^{n-1} p^{jk}[(\nu, p) = 1, 1 \leqq \nu \leqq p^{n-j}] \sum_{\nu} \nu^k,$$

(1, 12) follows from the fact mentioned in the introduction that

$$[(\nu, p) = 1, 1 \leqq \nu \leqq p^{n-j}] \sum_{\nu} \nu^k \equiv 0 \pmod{p^{n-j-1}}.$$

For a recent proof of this theorem, see [1]. Define t_k, k a positive integer, by $t_k(x) = x^k$. Then

$$\begin{aligned}
\sigma[i_0(p^e), p^n, t_k] &= [i \equiv i_0 \pmod{p^e}, 0 \leqq i \leqq p^n - 1] \sum_{i} (r_i)^k \\
&= [0 \leqq \nu \leqq p^{n-e} - 1] \sum_{\nu} (r_{i_0+\nu p^e})^k = \sum_{\nu} (r_{i_0} + r_\nu q^e)^k \\
&= \sum_{\nu} \sum_{\mu=0}^{k} C(k, \mu) (r_{i_0})^\mu \cdot (q^e)^{k-\mu} (r_\nu)^{k-\mu} \\
&= p^{n-e} r_{i_0}{}^k + \sum_{\mu=1}^{k-1} C(k, \mu) (r_{i_0})^\mu (q^e)^{k-\mu} \sum_{\nu} (r_\nu)^{k-\mu} \\
&\equiv p^{n-e} r_{i_0}{}^k + \sum_{\mu=1}^{k-1} C(k, \mu) (r_{i_0})^\mu (q^e)^{k-\mu} \sum_{\nu} (\nu)^{k-\mu} \pmod{p^{n-e}}.
\end{aligned}$$

by (1, 1), since the representatives are initially complete. An application of (1, 12) yields $\sigma[i_0(p^e), p^n, t_k] \equiv 0 \pmod{p^{n-e-1}}$, and hence $\delta[i_0(p^e), p^n, t_k] \leqq | p^{-e-1} |$. Now if g is given by the power series $\sum_{k=0}^{\infty} a_k x^k$, convergent for all p-adic integers, then $\lim_{k\to\infty} | a_k | = 0$. Let $| p^M | = \max_k | a_k |$. Then

$$| \delta[i_0(p^e), p^n, g] | = | \sum_{k=0}^{\infty} a_k \delta[i_0(p^e), p^n, t_k] | \leqq | p^M | \cdot | p^{-e-1} |.$$

Since the derivative of g is given by taking the derivative of the power series term by term, the same bound holds for the divided sums of $g^{(k)}$, for all k.

2. Evaluation of $\lim_{n\to\infty} D[i_0(p^e), p^n, g]$.

If $g(x)$ is given by a power series, the Bernoullian numbers may be introduced to evaluate $\lim_{n\to\infty} D[i_0(p^e), p^n, g]$ as an infinite series, the terms of which are products of a constant power of p, a sign factor, a Bernoullian number, and a coefficient of the power series of $g(x)$. $D[i_0(p^e), p^n, g]$ is a divided sum for the standard representatives in a residue class modulo p^e.

2, 1) THEOREM. *If $g(x)$ is a function defined for $| x - i_0 | \leqq | p^e |$ by the power series $g(i_0 + p^e y) = \sum_{k=0}^{\infty} a_k y^k$, convergent for all p-adic integers y, then*

$$(2,2)\qquad \lim_{n\to\infty} D[i_0(p^e), p^n, g] = p^{-e}[a_0 - a_1/2 + \sum_{\mu=1}^{\infty} a_{2\mu}(-1)^{\mu-1}B_\mu],$$

where B_μ is the μ-th Bernoullian number.

The power series development of $g(x)$ is used to transform $D[i_0(p^e), p^n, g]$ to a sum over all non-negative integers up to a given limit of positive powers. If $n \geqq e$,

$$D[i_0(p^e), p^n, g] = [i \equiv i_0 \pmod{p^e},\ 0 \leqq i \leqq p^n - 1] \sum_i g(i)/p^n$$

$$= [0 \leqq \nu \leqq p^{n-e} - 1] \sum_\nu g(i_0 + \nu p^e)/p^n = \sum_\nu \sum_{k=0}^{\infty} a_k \nu^k/p^n$$

$$= \sum_{k=0}^{\infty} a_k \left(\sum_\nu \nu^k/p^n \right).$$

In order to justify taking a term by term limit of the right member of this equation, it is sufficient to show that

$$(2,3)\qquad \lim_{n\to\infty} a_k \sum_\nu \nu^k/p^n \text{ exists for each } k,$$

$$(2,4)\qquad \sum_{k=0}^{\infty} a_k \left(\sum_\nu \nu^k/p^n \right) \text{ converges uniformly in } n.$$

For $k = 0$, $\lim_{n\to\infty} \sum_\nu \nu^0/p^n = p^{n-e}/p^n = p^{-e}$. For $k = 1$,

$$\lim_{n\to\infty} \sum_\nu \nu/p^n = \lim_{n\to\infty} (p^{n-e} - 1)p^{n-e}/2p^n = (p^{-e}/2) \lim_{n\to\infty} (p^{n-e} - 1) = -p^{-e}/2.$$

For $k \geqq 2$, by elementary number theory,

$$(2,5)$$

$$[0 \leqq \nu \leqq p^{n-e} - 1] \sum_\nu \nu^k/p^n = (p^{n-e})^{k+1}/(k+1) - (p^{n-e})^k/2 + [\rho \equiv 1 \pmod 2,$$

$$1 \leqq \rho \leqq k - 1] \sum_\rho (-1)^{(\rho-1)/2} B_{(\rho+1)/2} C(k, \rho)(p^{n-e})^{k-\rho}/(\rho+1),$$

where the $C(k, \rho)$ are binomial coefficients. Hence if k is even,

$$(2,6)$$

$$\lim_{n\to\infty} [0 \leqq \nu \leqq p^{n-e} - 1] \sum_\nu \nu^k/p^n$$

$$= \lim_{n\to\infty} (-1)^{[(k-1)-1]/2} B_{[(k-1)+1]/2} C(k, k-1)(p^{n-e})^{k-(k-1)}/p^n(k - 1 + 1)$$

$$= \lim_{n\to\infty} (-1)^{k/2-1} B_{k/2} k p^{n-e}/p^n k = (-1)^{k/2-1} p^{-e} B_{k/2},$$

and if k is odd,

$$(2,7)\qquad \lim_{n\to\infty} [0 \leqq \nu \leqq p^{n-e} - 1] \sum_\nu \nu^k/p^n = 0.$$

The limits for $k = 0, 1$ taken with (2, 6) and (2, 7), establish (2, 3). By (1, 12), $|a_k(\sum_p \nu^k/p^n)| \leqq |a_k|\,|p^{-e-1}|$, and (2, 4) follows. Since it is possible to take a term by term limit,

$$\lim_{n\to\infty} D[i_0(p^e), p^n, g] = \sum_{k=0}^{\infty} a_k \lim_{n\to\infty} (\sum_p \nu^k/p^n)$$

$$= p^{-e}[a_0 - a_1/2 + \sum_{\mu=1}^{\infty} a_{2\mu}(-1)^{\mu-1}B_\mu].$$

This proof might have been assembled from material in [4]. (2, 6) and (2, 7) can be obtained from reexamination of (2, 8) and (2, 10) of that paper. A term-by-term limit was taken to establish (4, 2) of that paper. Combination of (4, 2), (2, 8), and (2, 10) yields

$$\lim_{n\to\infty} [(i, p) = 1, 0 \leqq i \leqq p^n - 1] \sum_i f(i)/p^n$$

$$= a_0(1 - p^{-1}) + \sum_{j=1}^{\infty} a_{2j}(-1)^{j-1}B_j(1 - p^{2j-1}), \text{ where } f(x) = \sum_{j=0}^{\infty} a_j x^j.$$

The formula (2, 2) can be applied to $\exp(px) = \sum_{k=0}^{\infty} x^k(p^k/k!)$, convergent for all p-adic integers if $p \neq 2$, convergent if $x \equiv 0 \pmod 2$ for $p = 2$. (2, 2) gives

(2, 8) THEOREM. *If* $p \neq 2$, *then*

$$\lim_{n\to\infty} D[0(p), p^n, \exp] = p^{-1}[1 - p/2 + \sum_{\mu=1}^{\infty} (-1)^{\mu-1}B_\mu p^{2\mu}/(2\mu)!];$$

$$\lim_{n\to\infty} D[0(4), 2^n, \exp] = 2^{-2}[-1 + \sum_{\mu=1}^{\infty} (-1)^{\mu-1}B_\mu 2^{4\mu}/(2\mu)!].$$

The first limit can also be evaluated by elementary means, since

$$D[0(p), p^n, \exp] = [0 \leqq i \leqq p^{n-1} - 1] \sum_i \exp(ip)/p^n = (e^{p^n} - 1)/(e^p - 1)p^n,$$

by the formula for the sum of a geometric progression. Hence

$$\lim_{n\to\infty} D[0(p), p^n, \exp] = (e^p - 1)^{-1} \lim_{n\to\infty} (e^{p^n} - 1)/p^n = (e^p - 1)^{-1}, \tag{2, 9}$$

since $\lim_{n\to\infty} (e^{p^n} - 1)/p^n = (de^x/dx)_0 = 1$. Also

$$\lim_{n\to\infty} D[0(4), 2^n, \exp] = (e^4 - 1)^{-1}. \tag{2, 10}$$

The formal power series identity

$$x/(e^x - 1) = 1 - x/2 + \sum_{\mu=1}^{\infty} (-1)^{\mu-1}B_\mu x^{2\mu}/(2\mu)! \tag{2, 11}$$

is sometimes given (see [6], p. 127) as a definition of the Bernoullian numbers. Both members of (2, 11) are defined p-adically for $x = p$, p an odd prime, and for $x = 4$ when $p = 2$. Their equality follows from the two evaluations of $\lim_{n\to\infty} D[0(p), p^n, \exp]$.

For all p-adic integers α, the function $t_\alpha = (1 + px)^\alpha$ is defined and given by the binomial theorem $(1 + px)^\alpha = \sum_{k=0}^{\infty} C(\alpha, k)(px)^k$. An application of (2, 2) yields

(2, 12) THEOREM. *If α is a p-adic integer, then*

$$\lim_{n\to\infty} D[1(p), p^n, t_\alpha] = p^{-1}[1 - \alpha p/2 + \sum_{\mu=1}^{\infty} (-1)^{\mu-1} B_\mu p^{2\mu} C(\alpha, 2\mu)].$$

3. A p-adic γ function, Γ_p.

The function $\log x$, defined in residue class one modulo p, p odd, and in residue class one modulo 4, $p = 2$, by $\log(1 + py) = \Sigma(-1)^{k-1}(p^k/k)y^k$ affords another application of Theorem (2, 1).

(3, 1) THEOREM. *If p is an odd rational prime, then*

$$\lim_{n\to\infty} D[1(p), p^n, \log] = p^{-1}[-p/2 + \sum_{\mu=1}^{\infty} (-1)^{\mu} p^{2\mu} B_\mu/2\mu];$$

$$\lim_{n\to\infty} D[1(4), 2^n, \log] = 2^{-2}[-2 + \sum_{\mu=1}^{\infty} (-1)^{\mu} 2^{4\mu} B_\mu/2\mu].$$

The series of (3, 1) suggest the asymptotic Stirling series for the derivative of $\log \Gamma(x)$, where $\Gamma(x)$ is the classical gamma function.

If an attempt is made to evaluate $\lim_{n\to\infty} D[1(p), p^n, \log]$, p odd, by elementary means, a product analogous to the classical Γ-function for p-adic analysis arises. By definition,

$$\begin{aligned} D[1(p), p^n, \log] &= [i \equiv 1 \pmod p, 0 \leqq i \leqq p^n - 1] \sum_i \log i/p^n \\ &= [0 \leqq \nu \leqq p^{n-1} - 1] \sum_\nu \log(1 + \nu p)/p^n = \log \prod_\nu (1 + \nu p)/p^n \\ &= \log \Gamma_p(p^{n-1})/p^n, \end{aligned}$$

where $\Gamma_p(n)$ is defined for positive rational integers by

$$\Gamma_p(n) = \prod_{i=0}^{n-1} (1 + ip). \tag{3, 2}$$

Then

$$\lim_{n\to\infty} D[1(p), p^n, \log] = \lim_{n\to\infty} \log \Gamma_p(p^{n-1})/p^n = p^{-1} \lim_{n\to\infty} \log \Gamma_p(p^n)/p^n. \tag{3, 3}$$

For positive rational integers, $\Gamma_p(n)$ satisfies the difference equation

$$\Gamma_p(n+1) = (np+1)\Gamma_p(n) \tag{3,4}$$

with the initial condition that $\Gamma_p(1) = 1$. This difference equation is satisfied by

$$\Gamma_p(n) = [\Gamma(1+1/p)]^{-1} p^{n-1}\Gamma(n+1/p). \tag{3,5}$$

If the asymptotic series $\log x - 1/2x + \sum_{\mu=1}^{\infty} (-1)^{\mu} B_\mu / 2\mu x^{2\mu}$ for the derivative of $\log \Gamma(x)$ (see [3:101]) is used, the formula (3, 1) for p odd can be formally derived if $\lim_{n\to\infty} \log \Gamma_p(p^n)/p^n$ is interpreted as the derivative of $\log \Gamma_p(n)$ at zero.

For $p = 2$, the appropriate equations are

$$\Gamma_2(n) = \prod_{i=0}^{n-1} (1+4i), \qquad \lim_{n\to\infty} D[1(4), 2^n, \log] = 2^{-2} \lim_{n\to\infty} \log \Gamma_2(2^n)/2^n.$$

The difference equation $\Gamma_2(n+1) = (4n+1)\Gamma_2(n)$ with the initial condition $\Gamma_2(1) = 1$, is satisfied by $\Gamma_2(n) = [\Gamma(1+1/4)]^{-1} 4^{n-1}\Gamma(n+1/4)$.

The function $\Gamma_p(n)$ has been defined only for rational integers. It will be shown that the definition of $\Gamma_p(n)$ can be extended to all p-adic integers.

Suppose that n_1 and n_2 are positive integers with $n_1 \equiv n_2 \pmod{p^e}$, $n_1 \geqq n_2$. Write $n_1 = n_2 + kp^e$. Then

$$\Gamma_p(n_1) - \Gamma_p(n_2) = \Gamma_p(n_2)\{[n_2 \leqq i \leqq n_2 + kp^e - 1]\prod_i (1+pi) - 1\}$$

$$= \Gamma_p(n_2)\{\prod_{\rho=0}^{k-1} [n_2 + \rho p^e \leqq i \leqq n_2 + (\rho+1)p^e - 1]\prod_i (1+pi) - 1\}.$$

Since the p^e numbers $1+ip$, $i = n_2 + \rho p^e, \cdots, n_2 + (\rho+1)p^e - 1$ are incongruent modulo p^{e+1} and are all congruent to one modulo p, they are congruent in some order to the first p^e numbers in residue class one. Therefore

$$[n_2 + \rho p^e \leqq i \leqq n_2 + (\rho+1)p^e - 1]\prod_i (1+ip)$$

$$\equiv [0 \leqq i \leqq p^e - 1]\prod_i (1+ip)$$

$(\mathrm{mod}\ p^{e+1})$. The elements in residue class one form a subgroup of a cyclic group. Let a be a generating element. Then

$$[0 \leqq i \leqq p^e - 1]\prod_i (1+ip) \tag{3,6}$$

$$\equiv [0 \leqq i \leqq p^e - 1]\prod_i a^i \pmod{p^{e+1}} \equiv a^{(p^e-1)p^e/2} \equiv 1$$

$(\mathrm{mod}\ p^{e+1})$. Multiplication of the resulting congruences

$$[n_2 + \rho p^e \leqq i \leqq n_2 + (\rho+1)p^e - 1]\prod_i (1 + pi) \equiv 1 \ (\mathrm{mod}\ p^{e+1}),$$

$\rho = 0, \cdots, k-1$, gives the result that if $n_1 \equiv n_2 (\mathrm{mod}\ p^e)$, $n_1 \geqq n_2$, then

$$\Gamma_p(n_1) - \Gamma_p(n_2) = \Gamma_p(n_2)\Big[\prod_{i=n_2}^{n_1-1} (1 + ip) - 1\Big] \equiv 0 \ (\mathrm{mod}\ p^{e+1})$$

and $\Gamma_p(n_1) \equiv \Gamma_p(n_2) (\mathrm{mod}\ p^{e+1})$. Hence $\Gamma_p(n)$, p odd, satisfies a Lipschitz condition.

The discussion of the case $p = 2$ is similar. It suffices to show that

$$[0 \leqq i \leqq 2^e - 1]\prod_i (1 + 4i) \equiv 1 \ (\mathrm{mod}\ 2^{e+1}).$$

The residue classes of 2^{e+2} congruent to one modulo four form a cyclic group if $e \geqq 1$ (see [2], p. 49). The number five is a generating element. Hence

$$[0 \leqq i \leqq 2^e - 1]\prod_i (1 + 4i) \equiv [0 \leqq i \leqq 2^e - 1]\prod_i 5^i \equiv 5^{2^e(2^e-1)/2} \equiv 1$$

$(\mathrm{mod}\ 2^{e+1})$. Therefore $\Gamma_2(n)$ satisfies the same Lipschitz condition as $\Gamma_p(n)$, p odd. Only the weaker

(3, 6) THEOREM. *$\Gamma_p(n)$ is uniformly continuous*

will be required in the sequel. The considerations used in establishing (3, 6) are similar to the group theoretic proof of the generalized Wilson's theorem.

By a well-known topological theorem (see [7], p. 28), since $\Gamma_p(n)$ is defined and uniformly continuous on the rational integers, which are dense in the p-adic integers, Γ_p has a unique uniformly continuous extension to the p-adic integers, which will also be written as Γ_p, but with x as an argument. It will be established in the next section that $\log \Gamma_p(x)$ is analytic.

4. Continuity and analyticity of sum functions.

Let $g(x)$ be a function defined for all p-adic integers. Let us introduce as the sum function of g for generalized representatives the sum of the function values over the representatives of the integers up to n,

$$\sigma(n, g) = \sum_{i=0}^{n-1} g(r_i). \tag{4, 1}$$

For standard representatives write

$$S(n, g) = \sum_{i=0}^{n-1} g(i). \tag{4, 2}$$

Then $\sigma(n, g)$ satisfies the difference equation

$$\sigma(n+1, g) - \sigma(n, g) = g(r_n), \tag{4, 3}$$

with initial condition $\sigma(1, g) = g(r_0)$; and $S(n, g)$ is a solution of the difference equation

$$S(n+1, g) - S(n, g) = g(n), \tag{4, 4}$$

with initial condition $S(1, g) = g(0)$.

Given a particular solution of (4, 3) or (4, 4), the sum of this solution and a function of period one is again a solution. In the calculus of finite differences, a solution of (4, 4) is singled out by certain fairly complicated limit conditions (see [3], pp. 40-42). In the present discussion particular solutions are singled out by the requirement of uniform continuity. It will be shown that if $g(x)$ is uniformly continuous, the sum function $\sigma(n, g)$ is uniformly continuous.

Define the characteristic function of the neighborhood of i_0 with radius $|p^e|$ by

$$\begin{aligned} \mathrm{Ch}[x, i_0(p^e)] &= 1 \text{ if } x \equiv i_0 \pmod{p^e} \\ &= 0 \text{ if } x \not\equiv i_0 \pmod{p^e}. \end{aligned} \tag{4, 5}$$

Let the generalized representatives be initially complete, and let n_1 and n_2 be positive integers with $n_1 \equiv n_2 \pmod{p^m}$, or $n_1 = n_2 + kp^m$, where k is a positive integer. If $m \geqq e$, then

$$\begin{aligned} &\sigma(n_1, \mathrm{Ch}[x, i_0(p^e)]) - \sigma(n_2, \mathrm{Ch}[x, i_0(p^e)]) \\ &\quad = [n_2 \leqq i \leqq n_2 + kp^m - 1] \sum_i \mathrm{Ch}[r_i, i_0(p^e)] = kp^{m-e} \equiv 0 \pmod{p^{m-e}}, \end{aligned}$$

since $r_{i_1} \equiv r_{i_2} \pmod{p^\nu}$ if and only if $i_1 \equiv i_2 \pmod{p^\nu}$. This proves the

(4, 6) LEMMA. *If the generalized representatives are initially complete, then $\sigma(n, \mathrm{Ch}[i_0(p^e)])$ is uniformly continuous.*

The uniform continuity of the characteristic functions of neighborhoods will be used to prove

(4, 7) THEOREM. *If g is uniformly continous, and the generalized representatives are initially complete, then $\sigma(n, g)$ is uniformly continuous.*

Since g is uniformly continuous, for every positive integer ϵ there exists an N_1 such that if $i_1 \equiv i_2 \pmod{p^{N_1}}$, so that $r_{i_1} \equiv r_{i_2} \pmod{p^{N_1}}$, then $g(r_{i_1}) \equiv g(r_{i_2}) \pmod{p^\epsilon}$. Define g_ϵ by $g_\epsilon(x) = g(r_i)$ if $x \equiv r_i \pmod{p^{N_1}}$, $i = 0, \cdots, p^{N_1} - 1$. It follows that for all i,

$$g(r_i) \equiv g_\epsilon(r_i) \pmod{p^\epsilon}. \tag{4, 8}$$

By definition, g_ϵ is a linear combination of p^{N_1} characteristic functions; and hence, by (4, 6), $\sigma(n, g_\epsilon)$ is uniformly continuous. There exists an N_2 such

that if $n_1 \equiv n_2 \pmod{p^{N_2}}$, $n_1 \geqq n_2$, then $\sigma(n_1, g_\epsilon) \equiv \sigma(n_2, g_\epsilon) \pmod{p^\epsilon}$. But

$$\sigma(n_1, g_\epsilon) - \sigma(n_2, g_\epsilon) = \sum_{i=n_2}^{n_1-1} g_\epsilon(r_i) \equiv \sum_{i=n_2}^{n_1-1} g(r_i) \pmod{p^\epsilon} \text{ by } (4, 8),$$
$$\equiv \sigma(n_1, g) - \sigma(n_2, g) \pmod{p^\epsilon}.$$

Hence if $n_1 \equiv n_2 \pmod{p^{N_2}}$, then $\sigma(n_1, g) \equiv \sigma(n_2, g) \pmod{p^\epsilon}$. Thus the sum function $\sigma(n, g)$ is uniformly continuous, and consequently is capable of a unique uniformly continuous extension $\sigma(x, g)$ to the p-adic integers. This extended function is the p-adic solution of $(4, 3)$ singled out by the requirement of uniform continuity.

If $f(x)$ is analytic, and the coefficients of the power series satisfy certain additional restrictions, the sum function over standard representatives is analytic. Suppose that $f(x)$ is given by the power series

$$f(x) = \sum_{i=0}^{\infty} a_i x^i, \tag{4, 9}$$

convergent for all p-adic integers. The necessary and sufficient condition for convergence is that $\lim_{i \to \infty} | a_i | = 0$. Now

$$S(n, f) = \sum_{\mu=0}^{n-1} f(\mu) = \sum_{\mu=0}^{n-1} \left(\sum_{i=0}^{\infty} a_i \mu^i \right) = \sum_{i=0}^{\infty} a_i \left(\sum_{\mu=0}^{n-1} \mu^i \right),$$

and hence

$$S(n, f) = n a_0 + \sum_{i=1}^{\infty} a_i \{ n^{i+1}/(i+1) - n^i/2 + \sum_{\rho=1}^{[i/2]} (-1)^{\rho-1} B_\rho C(i, 2\rho - 1) n^{i-2\rho+1}/2\rho \}, \tag{4, 10}$$

where $[r]$ is the greatest integer $\leqq r$.

It must be shown that $(4, 10)$ can be rearranged to express $S(n, f)$ as a power series in n. The double series $\sum_i \sum_k b_{ik}$ can be summed in any order if the general term goes to zero, that is, if for every positive integer ϵ there exists an N such that if $i + k \geqq N$, then $| b_{ik} | \leqq | p^\epsilon |$. Let b_{ik} be the term of $(4, 10)$ containing a_{i-1} and n^k as factors. Then the $(m+1)$-st row of the double array of b_{ik} has at most the first $m+1$ elements not zero, since the elements of the $(m+1)$-st row are the terms of the summation

$$a_m \{ n^{m+1}/(m+1) - n^m/2 + \sum_{\rho=1}^{[m/2]} (-1)^{\rho-1} B_\rho C(m, 2\rho - 1) n^{m-2\rho+1}/2\rho \}.$$

Now $| C(m, 2\rho - 1) | \leqq 1$, and $| B_\rho | \leqq | p^{-1} |$ by the von Staudt-Clausen theorem of elementary number theory. Hence for all k,

$$(4,11)\qquad |b_{m+1\,k}| \leqq |p^{-1}| \max\{|a_m/(m+1)|, \max_{1\leqq\rho\leqq[m/2]} |a_m/2\rho|\}.$$

If for every positive integer ϵ there exists an N such that for all $m \geqq N$,

$$(4,12)\qquad |p^{-1}| \max\{|a_m/(m+1)|, \max_{1\leqq\rho\leqq[m/2]} |a_m/2\rho|\} \leqq |p^{\epsilon}|,$$

then if $i+k \geqq 2N+1$, $|b_{ik}| \leqq |p^{\epsilon}|$ by (4, 11) and (4, 12). If (4, 12) is satisfied by the coefficients of the power series of f, then (4, 10) can be summed in any order.

Rearrangement of (4, 10) gives

(4, 13)

$$S(n,f) = \sum_{\rho=1}^{\infty} n^{\rho}[a_{\rho-1}/\rho - a_{\rho}/2 + \sum_{k=1}^{\infty} (-1)^{k-1} a_{2k+\rho-1} B_k C(2k+\rho-1, 2k-1)/2k].$$

Write $S(x,f)$ for the uniformly continuous extension of $S(n,f)$ as given by the rearranged series (4, 13). Then $S(x,f)$ is given by the same power series with x as argument. The results of this discussion are summarized in

(4, 14) THEOREM. *If f is defined for all p-adic integers by the power series $f(x) = \sum_{i=0}^{\infty} a_i x^i$, and for every positive integer ϵ there exists an N such that for all $i \geqq N$,*

$$(4,15)\qquad \max\{|a_i/(i+1)|, \max_{1\leqq\rho\leqq[i/2]} |a_i/2\rho|\} \leqq |p^{\epsilon}|,$$

then $S(x,f)$ is an analytic function given by

$$(4,16)\qquad S(x,f) = \sum_{\rho=1}^{\infty} x^{\rho}[a_{\rho-1}/\rho - a_{\rho}/2 + \sum_{k=1}^{\infty} (-1)^{k-1} a_{2k+\rho-1} B_k C(2k+\rho-1, 2k-1)/2k].$$

Note that $\lim_{n\to\infty} S(p^n,f)/p^n$ is given by the first coefficient of the power series of $S(x,f)$. This remark can be translated into a weakened statement of Theorem (2, 1). It also permits the interpretation that $\lim_{n\to\infty} S(p^n,f)/p^n$, the limit of divided sums, is the derivative of the sum function $S(x,f)$ at zero.

The analyticity of $\log \Gamma_p$ is an easy consequence of Theorem (4, 14).

(4, 17) THEOREM. *If $p \neq 2$, $\log \Gamma_p(x)$ is given by the power series*

$$\log \Gamma_p(x) = \sum_{\rho=1}^{\infty} (-1)^{\rho} x^{\rho}\{p^{\rho-1}/\rho(\rho-1) + p^{\rho}/2\rho + \sum_{k=1}^{\infty} (-1)^{k-1} p^{2k+\rho-1} B_k C(2k+\rho-1, 2k-1)/(2k+\rho-1)2k\}.$$

Also

$$\log \Gamma_2(x) = \sum_{\rho=1}^{\infty} (-1)^{\rho} x^{\rho} \{4^{\rho-1}/\rho(\rho-1) + 4^{\rho}/2\rho + \sum_{k=1}^{\infty} (-1)^{k-1} 4^{2k+\rho-1} B_k C(2k+\rho-1, 2k-1)/(2k+\rho-1)2k\}.$$

For $p \neq 2$, $\log \Gamma_p(x) = S[x, \log(1+py)]$ by definition, and $\log(1+py)$ is given by the power series $\log(1+py) = \sum_{i=1}^{\infty} (-1)^{i-1}(py)^i/i$. It only remains to show that the coefficients $a_i = (-1)^{i-1}p^i/i$ satisfy the restriction (4, 15) on the coefficients of the power series. Since $|1/i| \leqq |p^{-i/p}|$, it follows that

$$\max\{|a_i/(i+1)|, \max_{1\leqq\rho\leqq[i/2]} |a_i/2\rho|\} = \max\{|p^i/i(i+1)|, \max_{1\leqq\rho\leqq[i/2]} |p^i/2i\rho|\}$$
$$\leqq \max\{|p^{i-(2i+1)/p}|, \max_{1\leqq\rho\leqq[i/2]} |p^{i-(i+\rho)/p}|\} = |p^{i-(2i+1)/p}|.$$

Hence (4, 15) is satisfied for $a_i = (-1)^{i-1}p^i/i$, if $i-(2i+1)/p \geqq \epsilon$, or $i \geqq (p\epsilon+1)/(p-2)$.

For $p=2$, $\log \Gamma_2(x) = S[x, \log(1+4y)]$ and $a_i = (-1)^{i-1}4^i/i$. Then

$$\max\{|a_i/(i+1)|, \max_{1\leqq\rho\leqq[i/2]} |a_i/2\rho|\} = \max\{|4^i/i(i+1)|, \max_{1\leqq\rho\leqq[i/2]} |4^i/2i\rho|\}$$
$$\leqq \max\{|2^{2i-(i+1)/2}|, \max_{1\leqq\rho\leqq[i/2]} |2^{2i-1+(i+\rho)/2}|\} \leqq |2^{2i-(2i+1)/2}| \leqq |2^{\epsilon}|,$$

if $2i-(2i+1)/2 \geqq \epsilon$, or $i \geqq \epsilon + 1/2$.

INDIANA UNIVERSITY.

REFERENCES.

[1] Hansraj Gupta, "An important congruence," *Proceedings of the Indian Academy of Sciences, Section A*, vol. 13 (1941), pp. 85-86.

[2] Erich Hecke, *Vorlesungen über der Theorie der algebraischen Zahlen*, Leipzig (1923).

[3] N. E. Nörlund, *Differenzenrechnung*, Berlin (1924).

[4] Gordon Overholtzer, "A new application of the Schur derivate," *Bulletin of the American Mathematical Society*, vol. 51 (1945), pp. 313-324.

[5] Issai Schur, "Ein Beitrag zur elementaren Zahlentheorie," *Sitzungberichte der Preussischen Akademie der Wissenschaften* (1933), pp. 145-151.

[6] E. T. Whittaker and G. N. Watson, *A course of modern analysis*, Cambridge (1940).

[7] G. T. Whyburn, *Analytic topology*, New York (1942).

A CLASS OF PARTIALLY ORDERED ABELIAN GROUPS RELATED TO KY FAN'S CHARACTERIZING SUBGROUPS.[1]

By A. H. CLIFFORD.

In a recent paper,[2] Ky Fan gives three simple conditions (I, II, III) on a partially ordered abelian group which are necessary and sufficient that it be (order and group) isomorphic with a subgroup G of the group $C(\Omega)$ of all continuous real-valued functions defined on a compact[3] space Ω with two properties P_1 and P_2 (stated in **1** below). This is Ky Fan's Theorem 1 (KF, p. 418). A subgroup G of $C(\Omega)$ is called by Ky Fan a "characterizing subgroup of $C(\Omega)$" if, in addition to P_1 and P_2, it also satisfies P_3 (see **1** below; these are respectively properties (5.1), (5.3), (5.2) in KF, p. 419). The reason for the term, and the importance of the concept, is brought out by Ky Fan's Theorem 3 (KF, p. 422): if two characterizing subgroups G, G^* of $C(\Omega)$, $C(\Omega^*)$, resp., are isomorphic, then the compact spaces Ω and Ω^* are homeomorphic.

In his Theorem 4 (KF, p. 424), Ky Fan shows that a fourth condition IV is sufficient, along with I, II, and III, that G be isomorphic with a characterizing subgroup of some $C(\Omega)$ with compact Ω. But Example 1 of **3** below shows that IV is not a necessary condition. In **2** below we state a weaker condition IV′ which is necessary,[4] but Example 2 of **3** shows that IV′ is not sufficient. Incidentally, these examples are all real linear spaces; Ky Fan's paper applies almost verbatim to partially ordered spaces. The author has been unable to find a natural condition on a partially ordered abelian group G, supplementary to I, II, III, which is both necessary and sufficient that G be isomorphic with a characterizing subgroup (or subspace) of some $C(\Omega)$ with compact Ω.

In his Theorem 1, Ky Fan takes for Ω the set Λ of maximal convex

[1] Received February 12, 1951.

[2] Ky Fan, "Partially ordered additive groups of continuous functions," *Annals of Mathematics*, vol. 51 (1950), pp. 409-427. This paper will be cited as "KF." The pertinent definitions and the conditions I-IV are repeated in **1** below.

[3] In accordance with KF (p. 409), a *compact space* shall mean a Hausdorff space with the property that every covering by open sets contains a finite subcovering.

[4] The necessity of IV′ is an easy consequence of Ky Fan's Lemma 5.3 (KF, p. 420). It is also stated explicitly in Lemma 7.2 (KF, p. 423), and proved assuming IV.

subgroups of G, defining the topology on Λ by the now classical method (stated in **1**) of Kakutani, Stone, and Gelfand. In his Theorem 4, he takes for Ω the set Σ of maximal singular subgroups of G, showing that Σ is a closed subset of Λ. Condition IV′ is simply that Σ be a subset of Λ. Σ is not in general closed, but, being a subspace of the compact space Λ, is completely regular. This fact is shown in **2** below (Theorem 2) without making use of the compactness of Λ. Regarded as an additive group of continuous real-valued functions on Σ, G satisfies six conditions, $P_1 - P_5$ and Q, stated in **1**. If Σ is compact, then (as Ky Fan shows) P_4, P_5, and Q are consequences of P_1, P_2, P_3, but otherwise they are completely independent thereof (examples in **3**). Groups satisfying these six conditions are characterizing of the underlying space Σ in the sense of Ky Fan's Theorem 3.

1. For the convenience of the reader, we repeat here some of Ky Fan's fundamental definitions (KF, pp. 410-411).

A *partially ordered abelian group* G is an abelian group, written additively, some of whose elements f are *non-negative* (written $f \geqq 0$) and in which (i) $0 \geqq 0$; (ii) $f \geqq 0$ and $-f \geqq 0$ imply $f = 0$; $f \geqq 0$ and $g \geqq 0$ imply $f + g \geqq 0$. $f \geqq g$ (or $g \leqq f$) means $f - g \geqq 0$. We shall denote by G^+ the set of non-negative elements of G.

An element f of G is called an *Archimedean element* of G if $f \geqq 0$ and if, for every $g \in G$, there exists a natural number n such that $nf \geqq g$. An element $f \geqq 0$ of G without this property is called a *non-Archimedean element* of G. A subgroup H of G is called a *singular subgroup* of G if (i) H contains no Archimedean element of G, and (ii) for any two elements f, g of H there is an h in H such that $h \geqq f$, $h \geqq g$. A subgroup H of G is a *convex subgroup* of G if $H \neq G$ and if $h_1 \in H$, $h_2 \in H$, $h_1 \geqq f \geqq h_2$ imply $f \in H$.

Ky Fan's conditions I, II, III, and IV are as follows.

I. *G contains a subgroup R which is isomorphic to the simply ordered additive group of all real numbers* (KF, p. 413).

II. *At least one element of R is an Archimedean element of G* (KF, p. 413).

III. *If $nf + g \geqq 0$ for all natural numbers n, then $f \geqq 0$* (KF, p. 417).

IV. *If a maximal convex subgroup N of G is not a maximal singular subgroup of G, then there is a finite set of elements $f_1, f_2, \cdots, f_n$ of G such that $f_i \notin N$ $(1 \leqq i \leqq n)$ and such that every maximal singular subgroup of G contains at least one of them* (KF, p. 423).

Let S be an abstract set, and let G be a set of real-valued functions defined on S. The Kakutani-Stone-Gelfand method of using G to define a topology in S is essentially the following. (It introduces just enough open sets in S to render the functions in G continuous.) *The neighborhoods of a point x_0 of S are subsets of S determined by x_0, any finite subset $f_1, \cdots, f_n$ of G, and any positive real number ϵ as follows: $U(x_0; f_1, \cdots, f_n; \epsilon)$ consists of all x in S satisfying* $\max_{1 \leqq i \leqq n} |f_i(x) - f_i(x_0)| < \epsilon$. We shall refer to this as the topology induced in S by G.

It will be convenient to list here a number of conditions applicable to any set G of real functions on a set S. We denote by G^+ the set of non-negative functions in G.

P_1. *G contains all constant functions.* (The unit function will be denoted by e: $e(x) = 1$ for all x in S.)

P_2. *If $x_1, x_2 \in S$ and $x_1 \neq x_2$, there exists $f \in G$ such that $f(x_1) \neq f(x_2)$.*

P_3. *If $f(x_0) = 0$ ($f \in G, x_0 \in S$), there exists $g \in G^+$ such that $g \geqq f$ and $g(x_0) = 0$.*

P_4. *If f is a non-Archimedean element of G then $f(x_0) = 0$ for some $x_0 \in S$.*

P_5. *If a set Φ of functions in G^+ has the property that any finite subset of Φ has a common zero in S then the whole set Φ has a common zero.*

We record also the following condition, applicable if S is a topological space.

Q. *If A is a closed subset of S, and $x_0 \notin A$, there exists a function $f \in G^+$ such that $f(x_0) = 0$ and* $\inf_{y \in A} f(y) > 0$.

THEOREM 1. *Let G be an additive group of real functions on S satisfying P_1, P_2, P_3.*

(1) *Relative to the topology of S induced by G, condition Q holds.*

(2) *For each $f \in G$ and each positive real number ϵ, the set $V(f; \epsilon)$ of all $x \in S$ such that $|f(x)| < \epsilon$ is an open subset of S. The set of all $V(f; \epsilon)$ with $f \in G^+$ and $\epsilon > 0$ is a basis for the open sets of S.*

(3) *If S is a topological space to begin with, the topology of S induced by G will coincide with its original topology if and only if, relative to the latter, the functions in G are continuous and condition Q holds.*

(4) *S is completely regular.*[5]

Proof. (1). Let A be a closed set in the topology of S induced by G, and let $x_0 \notin A$. Then there exists a neighborhood $U(x_0; f_1, \cdots, f_n; \epsilon)$ of x_0 having void intersection with A. Let $f_i(x_0) = \alpha_i$. By P_1, G contains $\alpha_i e$. Since G is an additive group, it contains the functions $f_i - \alpha_i e$, which vanish at x_0. By P_3, G^+ contains functions g_i vanishing at x_0 such that $g_i \geqq f_i - \alpha_i e$. Then the functions $h_i = g_i - f_i + \alpha_i e$ also belong to G^+. Let $f = g_1 + \cdots + g_n + h_1 + \cdots + h_n$. Then $f \in G^+$ and $f(x_0) = 0$. For any point y in A we have $y \notin U$, and hence $|f_i(y) - \alpha_i| \geqq \epsilon$ for some i.

$$f(y) \geqq g_i(y) + h_i(y) \geqq |g_i(y) - h_i(y)| = |f_i(y) - \alpha_i| \geqq \epsilon.$$

Hence $\inf_{y \in A} f(y) \geqq \epsilon > 0$.

(2). That $V(f; \epsilon)$ is open will follow when we show that to each $x_0 \in V$ there exists a neighborhood U of x_0 contained in V. Let $f(x_0) = \alpha$. Then $\eta = \epsilon - |\alpha| > 0$. As in (1), $g = f - \alpha e$ belongs to G. Let $U = U(x_0; g; \eta)$. If $x \in U$ then $|g(x)| < \eta$ and

$$|f(x)| = |g(x) + \alpha| \leqq |g(x)| + |\alpha| < \eta + |\alpha| = \epsilon$$

whence $x \in V$. To show that the sets $V(f; \epsilon)$ with $f \in G^+$ and $\epsilon > 0$ constitute a basis, we must show that if

$$x_0 \in W = V(f_1; \epsilon_1) \cap V(f_2; \epsilon_2) \qquad (f_1, f_2 \in G; \epsilon_1, \epsilon_2 > 0)$$

then there exist $f_3 \in G^+$ and $\epsilon_3 > 0$ such that $x_0 \in V(f_3; \epsilon_3) \subseteq W$. Let A be the complement of W in S. Since A is closed and $x_0 \notin A$, there exist by (1) $f_3 \in G^+$ and $\epsilon_3 > 0$ such that $\inf_{y \in A} f_3(y) \geqq \epsilon_3$, while $f(x_0) = 0$. If $x \in V(f_3; \epsilon_3)$ then $f_3(x) < \epsilon_3$, whence $x \notin A$ and $V(f_3; \epsilon_3) \subseteq W$.

(3). Assume that the functions in G are continuous and that Q holds in the original topology of S. Let N be an open subset of S in the original topology, and let $x_0 \in N$. Since Q holds, there exists $f \in G^+$ such that $f(x_0) = 0$, $f(x) \geqq \epsilon > 0$ for all x in the complement of N. Then $V(f; \epsilon) \subseteq N$, and N is open in the topology of S induced by G.

Conversely, we are to show that every subset of S open in the latter

---

[5] Part (4) is a consequence of two theorems in N. Bourbaki, *Topologie Générale* (Actualitées Scientifiques et Industrielles, nos. 858 and 1045). As a special case of a result on p. 117 of Chapter II, the space S is "uniformisable," while Theorem 2 on p. 9 of Chapter IX states that a space is uniformisable if and only if it is completely regular. The author is indebted to Professor Fan both for pointing out this reference and for supplying the quick proof of (4) which we give.

topology is also open in the original. By (2) it suffices to show this for every $V(f;\epsilon)$ with $f \in G^+$, $\epsilon > 0$. Let $x_0 \in V(f;\epsilon)$, i.e. $|f(x_0)| < \epsilon$. Let $\eta = \epsilon - |f(x_0)|$. Since f is continuous, by hypothesis, there exists a neighborhood N of x_0 such that $|f(x) - f(x_0)| < \eta$ if $x \in N$. Hence $x \in N$ implies

$$|f(x)| \leqq |f(x_0)| + |f(x) - f(x_0)| < (\epsilon - \eta) + \eta = \epsilon,$$

i.e. $x \in V(f;\epsilon)$, and hence $N \subseteq V(f;\epsilon)$.

(4). Let $U(x_0; f_1, \cdots, f_n; \epsilon)$ be a neighborhood of a point x_0 of S. We are to show the existence of a real continuous function g on S such that $g(x_0) = 0$, $g(x) = 1$ for $x \notin U$, and $0 \leqq g(x) \leqq 1$ for every $x \in S$. Such a function is the following:

$$g(x) = \min\{1, 1/\epsilon \max_{1 \leqq i \leqq n} |f_i(x) - f_i(x_0)|\}.$$

P_2 implies that every single-element subset of S is closed.

2. We now apply Theorem 1 to a partially ordered abelian group G satisfying Ky Fan's conditions I, II, III, and also:

IV′. *Every maximal singular subgroup of G is maximal convex.*

As in the proof of Ky Fan's Theorem 4 (KF, p. 424), G can be isomorphically represented as an additive group of real-valued functions on the set $\mathfrak{S}$ of maximal singular subgroups of G. The function corresponding to $f \in G$ is denoted by f. For $M \in \mathfrak{S}$, $\hat{f}(M)$ is the real number corresponding to the coset $f + M$ in the isomorphism between G/M and R.

Properties P_1 and P_2 are immediate for $S = \mathfrak{S}$, and P_3 is clear since each $M_0 \in \mathfrak{S}$ is singular, and $f(M_0) = 0$ is equivalent to $f \in M_0$. We may now topologize $\mathfrak{S}$ as in Theorem 1. $\mathfrak{S}$ becomes thereby a completely regular space; the functions f are bounded and continuous; and property Q holds.

THEOREM 2. *Any partially ordered abelian group G satisfying* I, II, III, IV′ *is isomorphic with an additive group of bounded continuous, real-valued functions on a completely regular space S (namely the space $\mathfrak{S}$ of maximal singular subgroups of G) satisfying $P_1 - P_5$ and Q.*

Conversely, let G be an additive group of bounded functions on a set S, and let G satisfy $P_1 - P_5$. Then G satisfies I, II, III, IV′. *There is a one-to-one correspondence between S and $\mathfrak{S}$ whereby the maximal singular subgroup M_{x_0} corresponding to the point x_0 of S consists of all $f \in G$ with $f(x_0) = 0$. If S is a topological space, if the functions in G are continuous, and if Q holds, then this correspondence is a homeomorphism.*

Proof. All of the first part has been shown except the validity of P_4 and P_5. P_4 is clear from the fact (KF, Lemma 2.2, p. 411) that every non-Archimedean element is contained in some maximal singular subgroup of G.

To show P_5, let Φ have the stated property, and let Ψ be the additive semigroup generated by Φ. If $f \in \Psi$ then $f = c_1 f_1 + \cdots + c_n f_n$ where each $f_i \in \Phi$ and each c_i is a positive integer. By the hypothesis on Φ, the functions f_i have a common zero M_0, i. e. $f_i \in M_0$. Then $f \in M_0$ also. Thus every element of Ψ is non-Archimedean, and the same holds for the set Ψ^* of all $f^* \in G^+$ satisfying $f^* \leqq f$ for some $f \in \Psi$. The difference group H of Ψ^* is therefore singular, and every such is contained in a maximal singular subgroup M_1 of G (KF, Lemma 2.2, p. 411). Since $\Phi \subseteq M_1$ we have $\hat{f}(M_1) = 0$ for every $f \in \Phi$, i. e. M_1 is a common zero of Φ.

Passing to the converse, let G be an additive group of bounded real functions on a set S, and let G satisfy $P_1 - P_5$. Conditions I, II, III are obvious. Before proving IV′, we shall establish the one-to-one correspondence $x_0 \leftrightarrow M_{x_0}$ between S and $\mathfrak{S}$. M_{x_0} is a singular subgroup of G by P_3, and the obvious fact that it contains no Archimedean element. It is evidently convex. In fact it is maximal convex. For if $f \notin M_{x_0}$, say $f(x_0) = \alpha \neq 0$, then $f - \alpha e \in M_{x_0}$. Any convex subgroup H of G containing M_{x_0} and f will therefore contain the Archimedean element $|\alpha| \cdot e$, and so will coincide with G. M_{x_0} is maximal singular. For any maximal singular subgroup of G containing M_{x_0} will be convex (KF, Lemma 2.4, p. 412), and hence will coincide with M_{x_0}. From P_2 it is clear that $x_1 \neq x_2$ implies $M_{x_1} \neq M_{x_2}$.

On the other hand, let M be any maximal singular subgroup of G, and let $M^+ = M \cap G^+$. Let $f_1, \cdots, f_n \in M^+$. Then $f_1 + \cdots + f_n = f \in M^+$. Since M contains no Archimedean elements, it follows from P_4 that $f(y_0) = 0$ for some $y_0 \in S$. From $f_1(y_0) + \cdots + f_n(y_0) = 0$ and $f_i(y_0) \geqq 0$ $(i = 1, \cdots, n)$ we conclude $f_i(y_0) = 0$. Thus any finite subset $f_1, \cdots, f_n$ of M^+ has a common zero in S. Applying P_5 we conclude the existence of $x_0 \in S$ such that $f(x_0) = 0$ for every $f \in M^+$. Since, by definition of singular subgroup, every element of M is the difference of two elements of M^+, we conclude that $M \subseteq M_{x_0}$. Then $M = M_{x_0}$ from the maximality of M and the singularity of M_{x_0}.

IV′ is now immediate since every maximal singular subgroup of G is some M_{x_0}, and the latter has been shown to be maximal convex. The last assertion of the theorem is immediate from conclusion (3) of Theorem 1.

We remark that if S is compact, then Ky Fan has shown that P_4, P_5,

and Q are consequences of P_1, P_2, P_3 and the continuity of the functions in G, G being then a characterizing subgroup of $C(S)$. P_4 follows from his Lemma 2. 2 and 5. 2. P_5 is evident from the compactness of S and the fact that the set of common zeros of a set of continuous functions is a closed set. Q is just Lemma 5. 5, together with the observation that f can be chosen to be positive. We conclude with a corollary which reduces to Ky Fan's Theorem 3 when S is compact.

COROLLARY. *Let G_1 and G_2 be additive groups of bounded continuous real-valued functions defined on topological spaces S_1 and S_2 resp., both groups satisfying $P_1 - P_5$ and Q. If G_1 and G_2 are isomorphic then S_1 and S_2 are homeomorphic. In particular, if one is compact, so is the other.*

Proof. By Theorem 2, S_i is homemorophic with the space $\mathfrak{S}_i$ of maximal singular subgroups of G_i $(i = 1, 2)$. But the topology in each $\mathfrak{S}_i$ is defined by purely algebraic means, so that the one-to-one correspondence between $\mathfrak{S}_1$ and $\mathfrak{S}_2$ induced in the obvious way by an isomorphism between G_1 and G_2 is a homeomorphism. We thus have $S_1 \sim \mathfrak{S}_1 \sim \mathfrak{S}_2 \sim S_2$.

3. In all of the following examples, S denotes the discrete space of the natural numbers $1, 2, \cdots$, and S^* the space S with the limit point ∞ adjoined. S is completely regular and S^* is compact. The additive group $BC(S)$ of bounded continuous real functions on S is just that of all bounded real sequences, and $BC(S^*) = C(S^*)$ is that of all convergent real sequences $\{f(n)\}$ with $f(\infty) = \lim_{n\to\infty} f(n)$.

Example 1 is a characterizing subgroup of $C(S^*)$ not satisfying Ky Fan's Condition IV (KF, p. 423). In fact it does not possess any finite subset such that every maximal singular subgroup contains at least one of them.

Example 2 is a subgroup G of $BC(S)$ satisfying $P_1 - P_5$ and Q. By Theorem 2 it must satisfy IV′. But from the Corollary thereto, since S is not compact, G cannot be isomorphic with a characterizing subgroup of $C(\Omega)$ for any compact Ω.

The remaining three examples show the independence of conditions P_4 and P_5 and Q under the assumption of P_1, P_2, P_3 and the continuity of the functions in G. To see the complete independence we may argue as follows.

Suppose the space S decomposes into two open-and-closed subsets S_1 and S_2. Then $BC(S)$ is the (cardinal) direct sum of $BC(S_1)$ and $BC(S_2)$. Let G_i be a subgroup of $BC(S_i)$, $i = 1, 2$, and let G be their direct sum; G is of

course a subgroup of $BC(S)$. If P is any one of the properties $P_1 - P_5$ or Q, then one readily sees that G has the property P if and only if both G_1 and G_2 have it. For example, the direct sum of Examples 3 and 4 gives an example in which P_4 and P_5 are false while Q is true. A simpler example for this case is the group of all convergent sequences (applied to S and not S^*). Likewise, for the case P_4 true, P_5 and Q both false, we may cite the group of all periodic sequences. But the author did not find any such easy examples for the two remaining cases.

Example 1. Let $a_1 = (1, 1/2, 1/3, 1/4, 1/5, \cdots)$, $a_2 = (1, 1, 1/2, 1/3, 1/4, \cdots)$, $a_3 = (1, 1, 1, 1/2, 1/3, \cdots), \cdots$ and as usual $e = (1, 1, 1, 1, 1, \cdots)$. Let G be the group of all sequences expressible as a finite linear combination

$$f = \lambda e + \rho_1 a_1 + \rho_2 a_2 + \cdots + \rho_n a_n \tag{1}$$

with real coefficients $\lambda, \rho_1, \rho_2, \cdots, \rho_n$. f is convergent with limit λ. We agree that $f(\infty) = \lim_{k\to\infty} f(k)$ for all $f \in G$ and G becomes thereby a subgroup of $C(S^*)$. We are to show that $P_1 - P_3$ hold, but IV does not. P_1 and P_2 are obvious.

To show P_3, let us first consider the point $x_0 = \infty$ of S^*. If f, as given by (1), vanishes at ∞ we must have $\lambda = 0$. But then

$$g = |\rho_1|\, a_1 + |\rho_2|\, a_2 + \cdots + |\rho_n|\, a_n$$

also vanishes at ∞ and has the properties $g \geqq 0$, $g \geqq f$. Now let $x_0 = k \in S$, and let $b_1 = 2e - 2a_1$, $b_k = 3e + 2a_{k-1} - 4a_k$ if $k > 1$. It is easily seen that $b_k(k) = 0$ while $b_k(j) \geqq 1$ for all $j \neq k$ (including $j = \infty$). If $f \in G$ and $f(k) = 0$, it is clear that some positive multiple of b_k will serve for g in P_3.

Hence P_3 hold, and G is a characterizing subgroup of $C(S^*)$. By Ky Fan's Lemma 5. 3, the maximal singular subgroups of G are simply the sets M_k of elements of G vanishing at $k \in S^*$. Consequently the functions $f_1, \cdots, f_n$ in Condition IV are to be sequences such that for each $k \in S^*$ there is at least one of the f_i such that $f_i(k) = 0$. The existence of such a set will be seen to be impossible when we show that no element $f \neq 0$ of G has more than a finite number of zeros in S.

Let f be given by (1), and suppose, by way of contradiction, that $f(k) = 0$ for infinitely many points k of S. Then surely $f(k_i) = 0$ for $n + 1$ points k_i satisfying $n < k_0 < k_1 < k_2 < \cdots < k_n$. Now for $k > n$,

$$f(k) = \lambda + \rho_1/k + \rho_2/(k-1) + \cdots + \rho_n/(k-n+1).$$

Then $f(k_i) = 0$ $(i = 0, 1, \cdots, n)$ is a system of $n + 1$ linear homogeneous

equations for $\lambda, \rho_1, \rho_2, \cdots, \rho_n$ with a determinant the value of which is readily calculated to be the ratio of

$$\prod_{i=1}^{n}(n-i)!\prod_{i>j}(k_i-k_j) \text{ to } \prod_{i=0}^{n}k_i(k_i-1)\cdots(k_i-n+1),$$

and therefore distinct from 0. Hence $\lambda=\rho_1=\rho_2=\cdots=\rho_n=0$. that is, $f=0$.

It might be objected that Condition IV could hold vacuously, in the sense that every maximal convex subgroup of G is also maximal singular. We show that this can not happen for any characterizing subgroup G of $C(\Omega)$ with a compact Ω consisting of more than one point. Let x_1, x_2 be distinct points of Ω, and define for each $f \varepsilon G$, $\tau(f)=\frac{1}{2}f(x_1)+\frac{1}{2}f(x_2)$. τ is a homomorphism of G into the simply ordered additive group of real numbers, and, by KF, Lemma 3.4, the set M of all $f \varepsilon G$ satisfying $\tau(f)=0$ is a maximal convex subgroup of G. Were M singular, every element f of M would be the difference $f=g-h$ of two positive elements g, h of M. But $g \geqq 0$ and $\tau(g)=0$ imply $g(x_1)=g(x_2)=0$, and similarly for h. Hence $f(x_1)=f(x_2)=0$. Thus M would be contained in the two distinct maximal convex subgroups M_{x_1} and M_{x_2} of G, contrary to its maximality.

Example 2. Let H be the group of all essentially finite sequences $(\alpha_1, \alpha_2, \cdots, \alpha_n, 0, 0, \cdots)$ with $\alpha_1+\alpha_2+\cdots+\alpha_n=0$. Evidently H contains no properly positive sequence. Let G consists of all sequences $\lambda e+h$ with λ any real number and $h \varepsilon H$. We proceed to show P_1-P_5 and Q, regarding G as a subgroup of $BC(S)$. P_1 and P_2 are evident.

Both P_3 and Q are evident when we exhibit, for each $k \varepsilon S$, an element b_k of G such that $b_k(k)=0$, $b_k(j) \geqq 1$ for all $j \neq k$. But such an element is $e+h_k$, where $h_k(j)=-1$ if $j=k$, $h_k(j)=1$ if $j=k+1$, and $h_k(j)=0$ for all other j.

To show P_4, suppose that $f=\lambda e+(\alpha_1, \alpha_2, \cdots \alpha_n, 0, 0, \cdots)$, $\sum_i \alpha_i=0$, is a non-Archimedean element of G. Were $f(k) \geqq 0$ for every $k \varepsilon S$ we would have $\lambda > 0$ and $\lambda+\alpha_i > 0$ for every $i=1, 2, \cdots, n$. This would contradict the condition $\inf_{k \varepsilon S} f(k)=0$ that f be non-Archimedean (KF, Lemma 5.1, p. 419).

To show P_5, let Φ be a set of positive functions of G with the stated property. Now a positive element $\lambda e+h$ of G must have $\lambda > 0$, and consequently can have only a finite number of zeros in S. Let $f_0 \varepsilon \Phi$, and let $k_1, \cdots, k_r$ be its zeros in S. If the conclusion to P_5 were false there would exist in Φ for each $i=1, \cdots, r$ an element f_i of G for which $f_i(k_i) \neq 0$.

But then the finite set $f_0, f_1, \cdots, f_r$ would have no common zero, contrary to hypothesis.

It is instructive to note that G can be regarded as a subgroup of $C(S^*)$, but fails to be a characterizing subgroup thereof since $M_\infty = H$ is not singular.

Example 3 (independence of P_4). Consider the group of Example 1 as a subgroup of $BC(S)$. Properties $P_1 - P_3$ are unaffected thereby. P_4 is false since a_1 is non-Archimedean but vanishes nowhere in S. P_5 follows from the fact that each element $\neq 0$ of G has only a finite number of zeros in S. Q follows from the existence of the sequences b_k.

Example 4 (independence of P_5). Let G be the group of all eventually constant sequences, regarded as a subgroup of $BC(S)$. $P_1 - P_4$ and Q are all evident. To show the falsity of P_5 one may take for Φ the following set of sequencies: $f_1 = (1, 0, 0, 0, \cdots)$, $f_2 = (1, 1, 0. 0, \cdots)$, $f_3 = (1, 1, 1, \cdots), \cdots$.

Example 5 (independence of Q). Let T denote the set S^* endowed with discrete topology. Let $G = C(S^*)$, but regarded as a subgroup of $BC(T)$. The properties $P_1 - P_5$ are unaffected thereby, but Q is false since there exists no $f \in G$ with $f(\infty) = 0$ and $\inf_{k<\infty} f(k) > 0$.

THE JOHNS HOPKINS UNIVERSITY.

A NOTE ON LIE ALGEBRAS OF CHARACTERISTIC p.*

By N. JACOBSON.

Let $\mathfrak{L}$ be a finite dimensional Lie algebra over a field of characteristic p, let $\mathfrak{A}$ be the universal associative algebra of $\mathfrak{L}$ ([1] and [4]) and let $\mathfrak{C}$ be the center of $\mathfrak{A}$. In this note we prove that if a is a linear element of $\mathfrak{A}$ then there exists a non-zero polynomial ϕ such that $\phi(a) \in \mathfrak{C}$. We use this result to obtain the following: (1) a simple direct proof of Iwasawa's theorem ([2], p. 420) that every finite dimensional Lie algebra of characteristic p has a faithful finite dimensional representation, (2) a proof of a conjecture of Chevalley that every finite dimensional Lie algebra of characteristic p has a representation which is not completely reducible, (3) a proof that $\mathfrak{A}$ can be imbedded in a division algebra.

1. A polynomial of the form $\alpha_0\lambda^{p^m} + \alpha_1\lambda^{p^{m-1}} + \cdots + \alpha_m\lambda$ is called a *p-polynomial.*

LEMMA. *Any polynomial $\neq 0$ is a factor of a suitable p-polynomial $\neq 0$.*

Proof. Let $\mu(\lambda) \neq 0$. For each $i = 0, 1, 2, \cdots$ write $\lambda^{p^i} = \mu(\lambda)q_i(\lambda) + r_i(\lambda)$ where $\deg r_i(\lambda) < \deg \mu(\lambda)$. Only a finite number of the $r_i(\lambda)$ are linearly independent. Hence we can find β_i not all 0 such that $\Sigma\beta_i r_i(\lambda) = 0$. Then $\Sigma\beta_i\lambda^{p^i} = \mu(\lambda)(\Sigma\beta_i q_i(\lambda))$.

PROPOSITION 1. *Let a be a linear element of the universal associative algebra $\mathfrak{A}$ of a finite dimensional Lie algebra $\mathfrak{L}$ of characteristic p. Then there exists a non-zero polynomial $\phi(\lambda)$ such that $\phi(a)$ is in the center $\mathfrak{C}$ of $\mathfrak{A}$.*

Proof. Let A denote the adjoint mapping $x \to [xa]$ in $\mathfrak{L}$. Since $\mathfrak{L}$ is finite dimensional there exists a non-zero polynomial $\mu(\lambda)$ such that $\mu(A) = 0$. Let $\phi(\lambda) = \Sigma\beta_i\lambda^{p^i}$ be a p-polynomial divisible by $\mu(\lambda)$. Then also $\phi(A) = 0$. For any $x \in \mathfrak{L}$ (regarded as the linear part of $\mathfrak{A}$)[1]

$$[x, \phi(a)] = \Sigma\beta_i[x, a^{p^i}] = \Sigma\beta_i[\cdots[x, \overbrace{a]\cdots a}^{p^i}] = x\Sigma\beta_i A^{p^i} = 0.$$

Hence $\phi(a) \in \mathfrak{C}$.

---

* Received April 28, 1951.

[1] We have used here the identity $[xa^p] = [\cdots[\overbrace{xa]\cdots a}^{p}]$. Cf. [3], p. 16.

Choose a basis $(e_1, e_2, \cdots, e_n)$ for $\mathfrak{L}$ and for each e_i let $\phi_i(\lambda)$ be a non-zero polynomial such that $y_i = \phi_i(e_i) \in \mathfrak{C}$. Let $d_i = \deg \phi_i(\lambda)$. Let $\mathfrak{B}$ be the ideal in $\mathfrak{A}$ generated by the elements y_i, $i = 1, 2, \cdots, n$.

PROPOSITION 2. *The cosets of the elements* $e_1^{\lambda_1} e_2^{\lambda_2} \cdots e_n^{\lambda_n}$, $0 \leq \lambda_i < d_i$, $(\lambda_1, \lambda_2, \cdots, \lambda_n) \neq 0$ *form a basis for* $\mathfrak{A}/\mathfrak{B}$.

Proof. It is known that the elements $e_1^{k_1} e_2^{k_2} \cdots e_n^{k_n}$, $k_i = 0, 1, 2, \cdots$, $(k_1, \cdots, k_n) \neq 0$ form a basis for $\mathfrak{A}$. Any polynomial in e_i can be expressed as a linear combination of elements $y_i^{m_i} e_i^{\lambda_i}$, $0 \leq \lambda_i < d_i$. Hence any element of $\mathfrak{A}$ is a linear combination of terms of the form $y_1^{m_1} y_2^{m_2} \cdots y_n^{m_n} e_1^{\lambda_1} e_2^{\lambda_2} \cdots e_n^{\lambda_n}$, $0 \leq \lambda_i < d_i$. It follows that every element is congruent mod $\mathfrak{B}$ to a linear combination of the elements $e_1^{\lambda_1} \cdots e_n^{\lambda_n}$. It remains to show that these elements are linearly independent mod $\mathfrak{B}$. Any element $b \in \mathfrak{B}$ is a linear combination of the terms $y_1^{m_1} y_2^{m_2} \cdots y_n^{m_n} e_1^{\lambda_1} e_2^{\lambda_2} \cdots e_n^{\lambda_n}$ where $(m_1, m_2, \cdots, m_n) \neq 0$. If we write

$$y_1^{m_1} \cdots y_n^{m_n} e_1^{\lambda_1} \cdots e_n^{\lambda_n} = y_1^{m_1} e_1^{\lambda_1} y_2^{m_2} e_2^{\lambda_2} \cdots$$

and express y_i in terms of e_i we obtain

$$y_1^{m_1} \cdots y_n^{m_n} e_1^{\lambda_1} \cdots e_n^{\lambda_n} = e_1^{\lambda_1 + m_1 d_1} e_2^{\lambda_2 + m_2 d_2} \cdots e_n^{\lambda_n + m_n d_n} + *$$

where the * is of degree $< \Sigma(\lambda_i + m_i d_i)$. It follows that if $b \neq 0$, then in the expression for b in terms of the basis $e_1^{k_1} e_2^{k_2} \cdots e_n^{k_n}$ at least one of the e_i has multiplicity $\geqq d_i$. Hence no non-zero linear combination of the terms $e_1^{\lambda_1} \cdots e_n^{\lambda_n}$, $\lambda_i < d_i$ belongs to $\mathfrak{B}$.

2. As a first application of Propositions 1 and 2 we prove

THEOREM 1. (*Iwasawa*) *Any finite dimensional Lie algebra of characteristic* $p \neq 0$ *has a faithful finite dimensional representation* ([2], p. 420).

Proof. Choose $y_i = \phi_i(e_i)$ of degree $d_i > 1$ in Proposition 2. Then the cosets $\bar{e}_i = e_i + \mathfrak{B}$ are linearly independent. Hence the natural imbedding of $\mathfrak{L}$ in $\mathfrak{A}/\mathfrak{B}$ is 1-1. On the other hand $\mathfrak{A}/\mathfrak{B}$ has the finite basis $e_1^{\lambda_1} e_2^{\lambda_2} \cdots e_n^{\lambda_n} + \mathfrak{B}$, $0 \leq \lambda_i < d_i$. Hence $\mathfrak{A}/\mathfrak{B}$ and, a fortiori, $\mathfrak{L}$ has a faithful finite dimensional representation.

THEOREM 2. (*Conjecture of Chevalley*) *Any finite dimensional Lie algebra of characteristic* p *has a finite dimensional representation which is not completely reducible.*

Proof. We now take $y_i = \phi_i(e_i)$ where $\phi_i(\lambda) = [\psi_i(\lambda)]^{n_i}$, $\deg \psi_i(\lambda) > 1$,

$\psi_i(e_i) \in \mathfrak{C}$ for all i, $n_i > 1$ for some i. Then $\bar{z}_i = \psi_i(e_i) + \mathfrak{B}$ is a non-zero element in the center of $\mathfrak{A}/\mathfrak{B}$. On the other hand, $\bar{z}_i^{n_i} = 0$ so that the center of $\mathfrak{A}/\mathfrak{B}$ has nilpotent elements. Hence the finite dimensional algebra $\mathfrak{A}/\mathfrak{B}$ is not semi-simple and therefore it has a finite dimensional representation which is not completely reducible. The same holds for $\mathfrak{L}$.

THEOREM 3. *The universal associative algebra of any finite dimensional Lie algebra of characteristic p can be imbedded in a division algebra.*

Proof. We can construct in the usual way the ring Δ of quotients ac^{-1}, $a \in \mathfrak{A}$, $c \neq 0$ in $\mathfrak{C}$. Δ is an integral domain and contains the quotient field Γ of $\mathfrak{C}$ in its center. If $(e_1, e_2, \cdots, e_n)$ is a basis for $\mathfrak{L}$ and $\phi_i(e_i) \in \mathfrak{C}$ where $\deg \phi_i = d_i > 0$ then we have seen that every element of $\mathfrak{A}$ is a linear combination with coefficients in $\mathfrak{C}$ (polynomials in $y_i = \phi_i(e_i)$) of the elements $e_1^{\lambda_1} e_2^{\lambda_2} \cdots e_n^{\lambda_n}$, $0 \leq \lambda_i < d_i$. It follows that Δ is finite dimensional over Γ. Hence Δ is a division algebra.

YALE UNIVERSITY.

REFERENCES.

[1] G. Birkhoff, "Representability of Lie algebras . . .," *Annals of Mathematics*, vol. 38 (1937), pp. 326-332.

[2] K. Iwasawa, "On the representation of Lie algebras," *Japanese Journal of Mathematics*, vol. 19 (1948), pp. 405-426.

[3] N. Jacobson, "Restricted Lie algebras of characteristic p," *Transactions of the American Mathematical Society*, vol. 50 (1941), pp. 15-25.

[4] E. Witt, "Treue Darstellung Liescher Ringe," *Journal für die reine und angewandedte Mathematik*, vol. 177 (1937), pp. 152-160.

ON THE LOGARITHMS OF BOUNDED MATRICES.*

By AUREL WINTNER.

Let capital letters denote (infinite) bounded matrices of complex numbers. Such a matrix, say A, is called non-singular if it has a unique (right- *and* left-sided) bounded reciprocal matrix. If the spectrum, $\mathrm{sp}\, A$, of a bounded A is defined to be the set of those complex numbers λ for which the matrix $\lambda I - A$ fails to be non-singular, then, as proved by me some time ago,[1]

(1) $\qquad \mathrm{sp}\, A \text{ is not vacuous}$

(*loc. cit.*, p. 243) and

(2) $\qquad \mathrm{sp}\, A \text{ is closed and bounded}$

(*loc. cit.*, p. 242), finally

(3) $\qquad \mathrm{sp}\, A \text{ is in } \{A\}$

(*loc. cit.*, p. 245). Here $\{A\}$ denotes the closure of the values attained by the bounded bilinear form, $A(x, y)$, of A under the following pair of restrictions: $y = \bar{x}$ and $|\, x\,| = 1$, where $\bar{x} = (\bar{x}_1, \bar{x}_2, \cdots)$ and $|\, x\,| = (|\, x_1\,|^2 + |\, x_2\,|^2 + \cdots)^{\frac{1}{2}}$, if $x = (x_1, x_2, \cdots)$. Corresponding to the Toeplitz-Hausdorff theorem on finite matrices,

(4) $\qquad \{A\} \text{ is convex.}$

As pointed out *loc. cit.* (pp. 246-249 and p. 244, respectively), neither (1) nor (3) can be concluded, by a limit process, from the corresponding facts on finite matrices, whereas (4) readily follows by such a process.

If $\mathbf{A}$ is any set of points in the complex plane, let $\mathbf{A}^*$ denote its complement. Then (2) and (1) show that $(\mathrm{sp}\, A)^*$ is an open set containing a domain $a < |\, \lambda\,| < \infty$ (if a is large enough) but not containing every point of the plane. Let $\mathbf{A}^0, \mathbf{A}^1, \cdots$ be the (finite or infinite) sequence representing the components of the open set $(\mathrm{sp}\, A)^*$, so that

(5) $\qquad \mathrm{sp}\, A = (\sum_n \mathbf{A}^n)^*,$

* Received June 28, 1951.

[1] A. Wintner, "Zur Theorie der beschränkten Bilinearformen," *Mathematische Zeitschrift*, vol. 30 (1929), pp. 228-282.

where every $\mathbf{A}^n$ is a connected open set which, if $m \neq n$, contains no point of $\mathbf{A}^m$. Let the notation be so chosen that $\mathbf{A}^0$ is the (unique) unbounded component of $(\mathrm{sp}\, A)^*$; so that

$$\sum_{n \neq 0} \mathbf{A}^n \text{ is bounded} \tag{6}$$

(possibly vacuous) and, according to (3),

$$\{A\}^* \text{ is in } \mathbf{A}^0. \tag{7}$$

The following considerations will contribute to the question of the existence of a bounded matrix function $f(A)$ of a given A. While the method will be presented in the important case $f(z) = \log z$ only, its applicability to more general functions $f(z)$ will be obvious. Incidentally, since it will consist of a process of analytic continuation, it will also apply to certain linear spaces more general than Hilbert's space.

(i) Corresponding to every component $\mathbf{A}^n$ of the complement $\sum \mathbf{A}^n$ of the spectrum of a bounded matrix A, there exists a bounded, non-singular matrix C_n which is commutable with A and has the following property: There belongs to every point λ of $\mathbf{A}^n$ a bounded matrix A_λ which is commutable with A and satisfies the relation

$$\lambda I - A = C_n \exp A_\lambda. \tag{8}$$

In addition, A_λ can be chosen so as to be continuous (and, for that matter, analytic and regular) on the λ-domain $\mathbf{A}^n$.

The latter notions refer to that metrization of the space of all bounded matrices A in which the norm $|A|$ is defined as the least upper bound of the vector length $|Ax|$ when x varies over all vectors of unit length.

Since $e^B e^C = e^{B+C}$ when B, C are bounded and commutable, it is clear that (i) contains the following corollary:

(ii) With reference to every fixed component $\mathbf{A}^n$ of $(\mathrm{sp}\, A)^*$, the matrix $\lambda I - A$ has a bounded logarithm either for every or for no λ contained in $\mathbf{A}^n$.

It is understood that a bounded matrix C is called a logarithm of a bounded matrix if the latter is e^C. Since $e^C e^{-C} = I = e^{-C} e^C$, only a non-singular matrix can have a bounded logarithm. If not every non-singular matrix can be represented as the square of a bounded matrix, then not every non-singular matrix has a bounded logarithm.

If $|Z|<1$, then $I+Z$ has a bounded logarithm. In fact, one such logarithm is supplied by the series

$$\log(I+Z)=\sum_{m=1}^{\infty} m^{-1}(-Z)^{m+1} \text{ if } |Z|<1. \tag{9}$$

This implies that if A is bounded and if the absolute value of λ is large enough, then $\lambda I-A$ has a bounded logarithm (in fact, (9) is applicable to $Z=-\lambda^{-1}A$ if $|\lambda|>|A|$). Since the matrix (9) is commutable with Z, it follows that (ii) contains the following corollary:

(iii) If λ is in $\mathfrak{A}^0$, i. e., in the unbounded component of the complement of sp A (e. g., in the complement of $\{A\}$), then $\lambda I-A$ has a bounded logarithm.

The parenthetical assertion of (iii) follows, by (3), from the main assertion of (iii).

If A is completely continuous, then sp A consists of the point $\lambda=0$ and of a finite (possibly vacuous) or infinite sequence of points which can cluster only at $\lambda=0$ (Hilbert). Since this implies that $(\text{sp } A)^*$ must be a connected set, and therefore identical with $\mathfrak{A}^0$, the first part of the following assertion is a consequence of (iii):

(iv) If A is completely continuous, then $\lambda I-A$ has a bounded logarithm for every λ not contained in sp A. In addition, this logarithm can always be chosen so as to be completely continuous.

The second part of (iv) will be clear from the proof of (i); cf. (II) below.

The proof of (i) centers about the following fact:

(I) If A and B are two bounded, commutable matrices the first of which is non-singular while the second is such as to satisfy

$$|B|<|A^{-1}|^{-1}, \tag{10}$$

then there exists a bounded matrix C which is commutable with A and with B and which satisfies the relation

$$A+B=Ae^{C}. \tag{11}$$

In fact, since $A+B$ is the product of A and $I+A^{-1}B=I+BA^{-1}$ (in either order), a C satisfying all the requirements of (I) results by chosing

$$Z=A^{-1}B \tag{12}$$

in (9). For, since (10) and (12) imply that $|Z| \leqq |A^{-1}|\,|B| < 1$, the proviso $|Z| < 1$ of (9) is satisfied.

Since the product of a bounded and of a completely continuous matrix is completely continuous, it is clear from this proof of (I), and from (12), that

(II) C in (I) can be chosen to be completely continuous if B is completely continuous.

It is easy to conclude from (I) the following lemma:

(III) Let A be a non-singular matrix corresponding to which there exists a family of non-singular matrices $A(t)$, $0 \leqq t \leqq 1$, satisfying the pair of conditions

$$A(0) = I, \qquad A(1) = A \tag{13}$$

and the commutability condition

$$A(u)A(v) = A(v)A(u). \tag{14}$$

Suppose further that

$$\text{g. l. b.}_{0 \leqq t \leqq 1} |A^{-1}(t)|^{-1} > 0 \tag{15}$$

and that $A(t)$ depends on t continuously. Then there exists a bounded matrix C satisfying $A = e^C$.

By the continuity of $A(t)$ is meant that, if ϵ is positive, and if $\delta = \delta_\epsilon > 0$ is small enough, then $|A(u) - A(v)| < \epsilon$ whenever $|u - v| < \delta$. Hence, if a is any positive constant, and if $j = j_a$ is large enough, then there exist $j+1$ values $0 = t_0 < t_1 < \cdots < t_j = 1$ satisfying

$$|A(t_k) - A(t_{k-1})| < a \text{ for } k = 1, \cdots, j. \tag{16}$$

Choose a to be the greatest lower bound (15). Then it is seen from (14) and (13), where $A(0) = \exp(0)$, that (III) follows by j consecutive applications of (I), namely, by identifying the matrices B and A in (I) with the matrices $A(t_k) - A(t_{k-1})$ and $A(t_{k-1})$, respectively.

It is readily seen that (i), the first of the assertions (i)-(iv) made above, is a corollary of (III). In fact, $\lambda I - A$ and $\mu I - A$ are always commutable, as are the reciprocals of $\lambda_1 I - A$ and $\lambda_2 I - A$ if λ_1 and λ_2 are not in sp A. On the other hand, since every component $\mathfrak{A}^n$, to which (i) refers, is an open connected set in the λ-plane, the existence of an appropriate continuous family $A(t)$, $0 \leqq t \leqq 1$, having the properties specified in (III), is obvious.

Remark. Suppose that A is a finite matrix. Then it has a logarithm, i. e., $A = e^C$ holds for some C, if (and only if) $\det A \neq 0$. In other words, $\lambda I - A$ has a logarithm whenever $\phi(\lambda) \neq 0$, where $\phi(\lambda) = \det(\lambda I - A)$. The standard proof of this theorem depends either on the existence of the Weierstrass-Jordan normal form of A, supplied by the theory of elementary divisors, or [2] (somewhat more directly, though in the main equivalently) on the Cayley-Hamilton theorem, $\phi(A) = 0$. The above considerations contain a proof which does not involve any such machinery. In fact, since sp A now consists of a finite number of points, the complement of sp A is $\mathfrak{A}^0$, hence the proof of (iii) or (iv) is applicable.

THE JOHNS HOPKINS UNIVERSITY.

[2] L. Schlesinger, "Über den Logarithmus einer Matrix," *Journal für die reine und angewandte Mathematik*, vol. 161 (1929), pp. 199-200. Cf. G. Pólya, "Der Picardsche Satz für Matrizen," *Jahresbericht der Deutschen Mathematiker Vereiningung*, vol. 40 (1931), second part, p. *80*.

ON PARALLEL SURFACES.*

By AUREL WINTNER.

1. Let $S: X = X(u, v)$, where $X = (x, y, z)$, be a surface of class C^2. Then, if S is sufficiently small, it can be assumed to be given in the form $z = z(x, y)$, where z has continuous partial derivatives of first and second order, $p = z_x, \cdots, t = z_{yy}$, and the total and mean curvatures, $K = K(x, y)$ and $H = H(x, y)$, of S are given by

$$Kd^4 = rt - s^2, \text{ where } d = (1 + p^2 + q^2)^{\frac{1}{2}}, \tag{1}$$

and

$$Hd^3 = \tfrac{1}{2}(1 + p^2)t - pqs + \tfrac{1}{2}(1 + q^2)r. \tag{2}$$

The sign of the first, but not of the second, of the continuous functions K, H has a geometrical meaning (H contains an odd power of the square root of the two-valued function $\pm d$).

Suppose that S is such as to make $K(x, y)$ a positive constant, say 1. Then (2) reduces to

$$z_{xx}z_{yy} - z_{xy}^2 = (1 + z_x^2 + z_y^2)^2, \tag{3}$$

a partial differential equation of second order which, according to S. Bernstein's general theorem on analytic differential equations of elliptic type, has the property of possessing no C^3-solution $z = z(x, y)$ which is not analytic as well. Accordingly, every surface of class C^3 having a constant positive curvature K is an analytic (for a more inclusive geometrical application of Bernstein's general theorem, cf. [1] and the comments in [6], Section 3).

This result contains, however, an unsatisfactory element which, along the lines of the preceding argument, would disappear only if it were known that Bernstein's general theorem (on arbitrary analytic differential equations of elliptic type in a plane) remains true for all solutions of class C^2. In fact those statements in differential geometry which require more than the mere existence of continuous derivatives of second order can hardly be justified for geometrical reasons. The restriction to the class C^3 (or, for that matter, $C^2(\lambda)$, where λ is a Hölder index; cf. [6], Section 5) often proves to be of an

* Received February 27, 1951.

artificial or preliminary nature, that is to say, such as to be necessitated by the lack of appropriate analytical tools only. Suffice it to observe that every surface of class C^2 has continuous first *and* second fundamental forms, and therefore, in particular, continuous curvatures. It is therefore of more than just formal interest that the above-mentioned consequence of Bernstein's general theorem can be freed of *all* of its non-geometrical restrictions, as follows:

(i) *If the total curvature of a surface of class C^2 is a positive constant, then the surface must be analytic.*

The point in this theorem is that, in contrast to the lemma of Section 3 (below) on which part of the proof of the theorem will depend, no data involving the second derivatives are subject to a Hölder restriction.

Needless to say, the analogue of the theorem is false if $K = \text{const.} \leqq 0$. This is obvious in the case const. $= 0$, since a torse can clearly be of class C^∞ without being analytic. Similarly, if const. $= -1 < 0$, then, since (1) becomes an analytic differential equation of hyperbolic type, it will surely possess non-analytic solutions $z = z(x, y)$ of class C^∞.

2. If $z(x, y)$ is a function possessing continuous second derivatives which are not subject to any restriction, then nothing but mere continuity can be claimed for the functions (1), (2). Similarly, if the second derivatives of $z(x, y)$ satisfy a Hölder condition (of some index λ), then the same is true (with the same λ) of both functions (1), (2). This makes clear the nature of the content of the following lemma:

(ii) *If a function $z = z(x, y)$ is of class C^2, and if its derivatives $p = z_x(x, y), \cdots, t = z_{yy}(x, y)$, besides being continuous, have the property that their combination $H = H(x, y)$, defined by* (2), *possesses continuous first derivatives H_x, H_y and that the latter satisfy, for some pair of positive constants γ, λ, a Hölder condition*

$$(4) \qquad |H_{x'} - H_{x''}| \leqq \gamma R^\lambda, \qquad |H_{y'} - H_{y''}| \leqq \gamma R^\lambda,$$

where $H_{x'} = H_x(x', y'), \cdots$ and $R^2 = (x' - x'')^2 + (y' - y'')^2$, then $z = z(x, y)$ must be of class C^3.

In particular, the function $K = K(x, y)$, being defined by (1), must then be of class C^1.

(ii bis) *The assertion of* (ii) *becomes false if $H(x, y)$ in the assumption* (4) *is replaced by $K(x, y)$.*

In order to see truth of this negation, (ii bis), let $z(x, y) = f(x + y)$, where $f(t)$ is a function of a single variable and has a continuous second derivative. Then $S: z = z(x, y)$ is of class C^2 and the function $K = K(x, y)$, being defined by (1), vanishes identically. Hence, (ii bis) follows from the circumstance that a function $f(t)$ possessing a continuous second derivative need not have a third derivative. (Nothing like this can, however, occur if the partial differential equation (1) for $z(x, y)$ is of elliptic type, that is, if $K(x, y) > 0$.)

3. The proof of (ii) proceeds as follows:

Let E_3 denote the three-dimensional Euclidean (z, p, q)-space, F_2 a (connected, open) domain in the (x, y)-plane, and L a function defined on the product space of E_3 and F_2 so as to satisfy the following conditions: If (x, y) is any fixed point of F_2, then L is fairly smooth (say, analytic) on E_3 and has a Hessian $L_{pp}L_{qq} - L_{pq}^2$ which is positive throughout; while, as a function of all five variables together, L is of class $C^1(\lambda)$, for some Hölder index $\lambda < 1$ (by virtue of the first of these conditions, the $C^1(\lambda)$-condition means that (4) is satisfied by $D = L$ and by a γ which can be chosen independent of (z, p, q) on every bounded portion of E_3). Under these assumptions, if $[L]$ denotes the Lagrangian derivative of L, every function $z = z(x, y)$ which is of class C^2 and satisfies the differential equation $[L] = 0$, on F_2, must be of class C^3. This follows from the proof, even though not from the wording, of a theorem of Lichtenstein [4]. In fact, although he assumes the analyticity of L in all five variables together, a glance at the proof shows what is actually needed is contained in the above conditions.

Let $H(x, y)$ denote a given function of class $C^1(\lambda)$ on an (x, y)-domain F_2. Then the conditions just required of L are satisfied by either of the functions

$$L(z, p, q; x, y) = d - H(x, y)z, \text{ where } d = \pm(1 + p^2 + q^2)^{\frac{1}{2}}.$$

But the Lagrangian equation $[L] = 0$ for a $z = z(x, y)$ then becomes precisely the condition (2), where $p = z_x, \cdots, t = z_{yy}$. Hence every function $z = z(x, y)$ of class C^2 which satisfies the relation (2), belonging to any given $H = H(x, y)$ of class $C^1(\lambda)$, must be a function of class C^3. This proves (ii).

4. For a given $n \geqq 1$, and for a sufficiently small piece, S, of a surface of class C^n, let

$$S: X = X(u^1, u^2), \tag{5}$$

where $X = (x, y, z)$, be a C^n-parametrization of S. By this is meant that, in the small (u^1, u^2)-domain under consideration, the (vector) function (5), besides being of class C^n, is such that the vector product $[X_1, X_2]$ of the partial derivatives $X_i = \partial X/\partial u^i$ does not vanish. Thus there exists on S a normal vector of unit length,

$$N = [X_1, X_2]/|[X_1, X_2]|, \tag{6}$$

which is a function $N(u^1, u^2)$ of class C^{n-1}. Hence, if c is a given number distinct from 0, then the (u^1, u^2)-parametrization of the X-set $S^* = S^*(c)$ which, in the sense of Steiner [5], is "the parallel surface of S at the distance c" (on one or on the other side of S, according as $c > 0$ or $c < 0$), that is, the parametrization

$$S^*: X = X^*(u^1, u^2) = X(u^1, u^2) + cN(u^1, u^2) \tag{7}$$

of the "parallel set" $S^* = S^*(c)$, is based on a function, $X^*(u^1, u^2)$, which is of class C^{n-1}. In fact, if the function (5) is of class C^n, then, in view of the differentiations contained in (6), the given function (7) will not in general be of class C^n. The following fact is therefore quite unexpected:

(iii) *If a surface S is of class C^2 and is sufficiently small with reference to a given real number c, then the corresponding parallel surface, $S^* = S^*(c)$, of S is again of class C^2, provided that* (with reference to the numerical values, K_0 and H_0, of the curvatures at that point of S the immediate vicinity of which is considered) *two numerical values of c are excepted*; *cf.* (19) *below.*

This does not mean, of course, that, barring the excepted c-values, (7) is a local C^2-parametrization whenever (5) is, but that (7) must then possess *some* local C^2-parametrizations,

$$S^*: X = X^*(v^1; v^2), \tag{7 bis}$$

in terms of *certain* parameters v^1, v^2 (which, in general, must be distinct from the given u^1, u^2). On the other hand, *both* $z(x, y)$ and $z^*(x, y)$ in

$$S: z = z(x, y) \quad \text{and} \quad S^*: z = z^*(x, y) \tag{8}$$

must be functions of class C^2 in some circle $x^2 + y^2 < b = b(a; c)$ if the local C^2-parametrization (5) of S is replaced by the Cartesian representation $z = z(x, y)$ in a circle $x^2 + y^2 < a$ the plane of which is the plane tangent to S at $(x, y) = (0, 0)$. In fact, the normal of S at $(0, 0)$ then becomes the z-axis, and the same is true of the normal of S^* at $(0, 0)$. It follows therefore from (iii), and from a standard application of the C^2-form of the

classical theorem on implicit functions, that both functions (8) must be of class C^2 in a vicinity of $(x, y) = (0, 0)$.

5. The actual content of (iii) is revealed by the second part ($n = 1$) of the following assertion:

(iii*) *The C^2 in* (iii) *can be replaced by C^n if $n \geqq 2$ but not if $n = 1$.*

While the first assertion of (iii*) will be clear from the proof (iii) below, failure for $n = 1$ can be concluded from the following example: For small $|x|$, let $f(x)$ be a function possessing a continuous derivative, $f'(x)$, and let $f(0) = 0$ and $f'(0) = 0$. Then, if $z(x, y)$ is defined to be $f(x)$ for all y, the surface S: $z = z(x, y)$ is of class C^1 and, since $f'(0) = 0$, the plane $z = 0$ is tangent to S along the y-axis. Hence, if a parallel surface $S^* = S^*(c)$ of $S: z = f(x)$ is of class C^1, then it must be representable, for small $|x|$, in the form $z^* = f^*(x)$, where $f^*(x)$ has a continuous first derivative which vanishes at $x = 0$. But it turns out that such is not the case if $f(x)$ is suitably chosen.

First, since S and S^* are represented by $z = f(x)$ and $z^* = f^*(x)$, it is seen from (6) and (7), where $X = (x, y, z) = (u^1, u^2, f)$ and $X^* = (u^1, u^2, f^*)$, that the function $z^* = f^*(x)$ is the result of the elimination of x between the two equations

$$x = u - f'(u)(1 + f'^2(u))^{-\frac{1}{2}}, \qquad z^* = f(u) + (1 + f'^2(u))^{-\frac{1}{2}},$$

if, without loss of generality, $c = 1$ in $S^* = S^*(c)$. Here $f'(u)$ is any continuous function vanishing at $u = 0$, and $f(u)$ is its primitive vanishing at $u = 0$. Accordingly, if $f' = g$, then for small $|u|$,

$$x = u - g(u)(1 + g^2(u))^{-\frac{1}{2}}, \tag{9}$$

where $g(u)$ is any continuous function satisfying $g(0) = 0$, and

$$z^* = (1 + g^2(u))^{-\frac{1}{2}} + \int_0^u g(v)\,dv. \tag{10}$$

The assertion is that, if $g(u)$ is suitably chosen, and if $z^* = f^*(x)$ denotes, for small $|x|$, the (continuous) function defined by the parametric equations (9)-(10), then $f^*(x)$ can fail to have a continuous derivative which vanishes at $x = 0$.

The square root, $(\ )^{-\frac{1}{2}}$, in (9)-(10) must be chosen positive if, corresponding to the choice $c = 1$, the value of the ordinate $z^* = f^*(x)$ at $x = 0$

is normalized to be $+1$. Then, since $(1+g^2)^{-\frac{1}{2}} = 1 - \frac{1}{2}g^2 + o(g^2)$ as $u \to \pm 0$, it follows that, in view of (10), the function $z^* = f^*(x)$ must satisfy the relation

$$f^*(x) - f^*(0) = -\tfrac{1}{2}g^2(u) + o(g^2(u)) + \int_0^u g(v)dv \tag{11}$$

by virtue of the relation (9), and that the latter is of the form

$$x = u - g(u) + o(|g(u)|^3). \tag{12}$$

Here $g(u)$ can be chosen to be any continuous function which vanishes at $u = 0$. But it is not hard to construct a (sufficiently "wobbly") continuous function $g(u) = o(1)$ which has the property that the ratio of (11) to (12) fails to tend to a limit as $u \to 0$ (even if u-values at which the denominator, (12), vanishes, are disregarded; these u-values can be confined to a sequence which clusters only at $u = 0$). For such a $g(u)$, the ratio of $f^*(x) - f^*(0)$ to x fails to tend to a limit as $x \to 0$. Since this means that $f^*(x)$ is non-differentiable at $x = 0$, the truth of the negative assertion of (iii*) follows.

6. In order to prove (iii), suppose that S is of class C^2 (and that (5) is a C^2-parametrization of S). Then S, besides having a first fundamental form $g_{\alpha\beta}du^\alpha du^\beta$ in which the coefficients

$$g_{ik} = X_i \cdot X_k = g_{ki} \qquad (\det g_{ik} > 0) \tag{13}$$

are functions of class C^1 in (u^1, u^2), possesses a second fundamental form $h_{\alpha\beta}du^\alpha du^\beta$ in which the coefficients

$$h_{ik} = -X_i \cdot N_k = h_{ki} \qquad (h_{ik} = N \cdot X_{ik}) \tag{14}$$

are continuous (possibly non-differentiable) functions of (u^1, u^2), the vector function (6) being of class C^1. The first derivatives of (6) are connected with those of (5) by Weingarten's identities

$$N_i = -h_{i\alpha}g^{\alpha\beta}X_\beta, \tag{15}$$

where, in matrix notations, $(g^{ik}) = (g_{ik})^{-1}$. In terms of (13) and (14), the representations of the total and mean curvatures, K and H, are

$$K = (\det h_{ik})/(\det g_{ik}), \qquad H = \tfrac{1}{2}g^{\alpha\beta}h_{\alpha\beta}. \tag{16}$$

With reference to any real number c, define the function $X^*(u^1, u^2)$ by (7). This vector function is of class C^1, since (6) is. It is readily verified

from (13)-(16) and (6) that the vector product of the partial derivatives $X^*_i(u^1, u^2)$ satisfies the identity

$$[X^*_1, X^*_2] = Q[X_1, X_2], \tag{17}$$

where $Q = Q^c = Q^c(u^1, u^2)$ is the (continuous, scalar) function

$$Q = 1 - 2cH + c^2K. \tag{18}$$

Let K_0, H_0 denote the values of K, H at any given point, $(u^1{}_0, u^2{}_0)$, of the (u^1, u^2)-domain and let, with reference to this point, the value of the arbitrary parameter c be chosen so as to be distinct from the roots of the quadratic equation

$$1 - 2cH_0 + c^2K_0 = 0 \tag{19}$$

(which becomes a linear equation or no equation at all according as $(u^1{}_0, u^2{}_0)$ is a non-umbilical or umbilical parabolic point of S, i. e., according as $K_0 = 0 \neq H_0$ or $K_0 = 0 = H_0$). Then the function (18), being distinct from 0 at the point $(u^1{}_0, u^2{}_0)$ of S, will be distinct from 0 in a vincinity of this point.

Let (u^1, u^2) be confined to such a vicinity. Then (17), where $[X_1, X_2] \neq 0$ by assumption, implies that $[X_1{}^*, X_2{}^*] \neq 0$. This means that (7) defines a surface $S^* = S^*(c)$ of class C^1. But (17), where $Q \neq 0$, also shows that the unit normal vector at a point (u^1, u^2) of S^* is identical with the unit normal vector, (6), of the S at the corresponding point (u^1, u^2) of S (with the understanding that, since (6) contains the ambiguity of a square root, N and $-N$ are not considered as distinct). Since (6) is a function of class C^1, it follows that S^* has a unit normal which is a function of class C^1 in the (u^1, u^2)-parametrization of S^*. It follows therefore by an application of the implicit function theorem that, if the plane tangent to S^* at $(u^1{}_0, u^2{}_0)$ is chosen to be the (x, y)-plane of the space $X = (x, y, z)$, then a (sufficiently small) vicinity of the point $(u^1{}_0, u^2{}_0)$ of S^* can be represented in the form $z = z(x, y)$, where $z(x, y)$ is a function of class C^2 (for details, cf. the corresponding argument in [3], p. 163). This proves (iii).

7. Denote by g^*_{ik}, h^*_{ik}, K^*, H^* the functions which result if $X(u^1, u^2)$ is replaced by $X^*(u^1, u^2)$ in the definitions (13), (14), (16), and refer by (j*), where $j = 6, 13, \cdots$, to the formula which, on this replacement, results from formula (j). These definitions of the starred functions are valid even though $X^*(u^1, u^2)$ need not be of class C^2 in the (u^1, u^2)-parameters. However, it is clear from the tensor character of the matrices of the first

and second fundamental forms that K^* and H^* are actually the total and mean curvatures, respectively, of S^*, and that (g^*_{ik}) and (h^*_{ik}) are the matrices which result by a calculation of the first and second fundamental forms in a C^2-parametrization and then applying the standard transformation rule.[1]

It is readily verified from (15) and (13) that $N_i \cdot N_k = g^{\alpha\beta} h_{i\alpha} h_{k\beta}$. In terms of matrix multiplications, this can be written in the form

$$(N_i \cdot N_k) = (h_{ik})\,(g^{ik})\,(h_{ik}) \tag{20}$$

(incidentally, if $\alpha = (g^{ik})$ and $\beta = (h_{ik})$, and if a prime denotes transposition, then the matrix product (20) can be written as $\alpha\beta\alpha'$, since $\alpha' = \alpha$). Since $N^* = N$, it follows from (20) that

$$(h^*_{ik})\,(g^{*ik})\,(h^*_{ik}) = (h_{ik})\,(g^{ik})\,(h_{ik}). \tag{21}$$

The connection between the respective factors in (21) can be obtained as follows:

Since (7) implies that $X^*_i = X_i + cN_i$, it is seen from (13) and (13*) that g^*_{ik} is the sum of three terms, g_{ik}, $cX_i \cdot N_k + cX_k \cdot N_i$ and $c^2 N_i \cdot N_k$. In view of (14) and (15), this means that

$$(g^*_{ik}) = (g_{ik}) + 2c(h_{ik}) + c^2 (h_{ik})\,(g^{ik})\,(h_{ik}) \tag{22}$$

and it is similarly verified, from (14) and (14*), that

$$(h^*_{ik}) = (h_{ik}) - c(h_{ik})\,(g^{ik})\,(h_{ik}). \tag{23}$$

Needless to say, the identity (21) could be verified from the relations (22) and (23). If the latter are combined with (16), (16*) and (18), there result the relations

$$K = QK^*, \qquad H = Q(H^* + cK^*). \tag{24}$$

While (24) is classical for a "smooth" pair S, S^* (in fact, for an S of class C^3, the corresponding S^* being of class C^2 for trivial reasons), the point in the following proof of (i) will be that (24) holds under the mere C^2-assumption of (iii).

8. Let S be a surface of C^2 and of constant positive curavture K.

---

[1] Needless to say, the matrix of the second fundamental form behaves under (u^1, u^2)-transformations (of class C^2) as a tensor only if the Jacobian of the transformation is positive, acquiring the factor -1 if the Jacobian is negative. This is precisely the reason why, in contrast to the case of the total curvature K, only the magnitude, and not the sign, of the mean curvature H has geometrical significance.

Then, with reference to any point $(u^1{}_0, u^2{}_0)$ of S, the quadratic equation (19) reduces to $Q_0 = 0$, where $Q_0 = Q(u^1{}_0, u^2{}_0; c)$ and, according to (18),

$$Q = Q(u^1, u^2; c) = 1 - 2cH + c^2, \qquad H = H(u^1, u^2),$$

if the value of the positive constant K is chosen to be 1. Since the product of the two roots of $Q_0 = 0$ is 1, either $c = 1$ or $c = -1$ is not a root (one of these c-values is a root if and only if the square of H_0 is 1, which, since $K_0 = K = 1$, means that $(u^1{}_0, u^2{}_0)$ is an umbilical point of S). It can be assumed that $c = 1$ is not a root, since, in view of the remark made after formula (2) above, H can be replaced by $-H$. Then $c = 1$ is not excepted in (iii), and so $S^* = S^*(1)$ is of class C^2.

The assumptions, $K = 1$ and $c = 1$, reduce (18) to $\frac{1}{2}Q = 1 - H$ (and $H = H(u^1, u^2)$ is distinct from 1 on S if S is a sufficiently small vicinity of its point $(u^1{}_0, u^2{}_0)$, since $Q_0 \neq 0$). Thus (24) reduces to the identities

$$1 = 2(1 - H)K^*, \qquad H = 2(1 - H)(H^* + K^*).$$

But substitution of the second of these identities into the first gives

$$H = 2(1 - H)H^* + 1, \text{ i. e., } (1 - H)(2H^* + 1) = 0$$

and (since $H \neq 1$ by the preceding parenthetical remark) the last equation implies that the function H^* of (u^1, u^2) must be the constant $-\frac{1}{2}$. This is the extension to the present case (of an S which is just of class C^2) of a classical result of Bonnet (cf., e. g., [2], p. 447).

9. The truth of (i) can now be concluded from (ii).

First, if S and H in (ii) are replaced by S^* and H^*, then, since S^* is of class C^2 and since H^*, being a constant, surely is of class $C^1(\lambda)$, it follows that S^* is of class C^3. Consequently, if S^* is written in the form $z = z^*(x, y)$ (locally), then $z^*(x, y)$ is a function of class C^3. Since $z^*(x, y)$ satisfies the differential equation which results if $H = H(x, y)$ in (2) is replaced by the constant $H^* = \frac{1}{2}$, it now follows from the general theorem of S. Bernstein (concerning elliptic differential equations; cf. Section 1) that the surface S^*: $z^* = z^*(x, y)$ is analytic. But the given S is a parallel surface of this S^*. Hence, the analyticity of S follows from that of S^*. This completes the proof of (i).

10. A corollary is the possibility of reducing to a geometrically significant minimum the smoothness assumptions in Liebmann's theorem concerning spheres, as follows:

If a closed, orientable surface is of (genus 0 and of) *class* C^2, *then* $K \equiv 1$ *cannot hold on it unless it is a sphere* (where, according to Hadamard [or just Gauss-Bonnet], the parenthetical restriction is superfluous).

Because of the use lines of curvatures in his partial differential equations, Hilbert's proof, as it stands, applies only under a C^n-restriction of fairly large n (something like $n = 5$); in this regard, cf. [3], pp. 163-172. But (i) reduces this n to $n = 2$ (S. Bernstein's reduction reaches to $n = 3$).

Appendix.

On tangential representations.

Let the surface (5) or $S\colon z = z(x, y)$ be of class C^2, sufficiently small and free of parabolic points (i. e., let either $K > 0$ or $K < 0$ on S). Then N, being defined by (6), is a vector function of class C^1, and the (local) theorem on implicit functions shows that the correspondance between $S\colon X = (x, y, z)$ and its normal image $N\colon (\xi, \eta, \zeta)$ is one-to-one and such that

$$S\colon X = X(N) \tag{25}$$

is a vector function of class C^1 (on the corresponding portion of the unit sphere $|N| = 1$). Hence the same is true of the scalar function $p = N \cdot X$ (Weingarten; cf. [2], pp. 135-137) and therefore of its homogenized form [2]

$$\begin{gathered} H = \rho p(\xi/\rho, \eta/\rho, \zeta/\rho), \text{ where } \rho = (\xi^2 + \eta^2 + \zeta^2)^{\frac{1}{2}} > 0, \\ N = (\xi/\rho, \eta/\rho, \zeta/\rho). \end{gathered} \tag{26}$$

Thus, if $S\colon X(u^1, u^2)$ is of class C^2, sufficiently small and of non-vanishing $K = K(u^1, u^2)$, then $H = H(\xi, \eta, \zeta)$ is a function of class C^1 (within that infinite cone of the (ξ, η, ζ)-space which corresponds to S in the spherical mapping of Gauss). But it turns out that, under the assumptions just stated, *the function* $H = H(\xi, \eta, \zeta)$, *instead of being only of class* C^1, *is of class* C^2 *by necessity.*

This lemma, which in a certain sense is an analogue of (iii) in Section 4 above, is unexpected, since it implies that, if parabolic points are excluded,

---

[2] This H, which is Minkowski's notation for the supporting function (26) (if $K < 0$ is excluded and S is closed and orientable), will not be confused with the preceding H, which is the mean curvature.

then Minkowski's applications of his *general* theory to *differential* geometry hold under the true assumption of the differential geometry of surfaces, which is the assumption of surfaces of class C^2. In contrast, the literature consulted says or implies that Minkowski's theory is applicable only under the artificial assumption that the surfaces considered are confined to the class C^3. Suffice it to say that the mean curvature and the Gaussian curvature are expressible in terms of the second derivatives of H provided that H is of class C^2, while the argument applied after (25) assures the C^2-character of H (i. e., of the representation of S in *plane coordinates*) only if (5) (i. e., the representation of S in terms of *point coordinates*) is of class C^3.

Incidentally, it will be clear from the proof below that the last italicized assertion can be generalized, to every $n \geqq 2$, in the same way as (iii*) in Section 5 generalizes (iii) in Section 4. In the excluded case, $n = 1$, the assertion is now not only false but meaningless as well, since the curvature $K = K(u^1, u^2)$ of a surface of class C^n cannot in general be defined if $n = 1$. On the other hand, it is clear (even in the case of an analytic S) that the assumption $K \neq 0$ cannot be omitted.

The proof (for $n = 2$) of the last italicized assertion proceeds as follows:

Since the piece of surface (5) is supposed to be sufficiently small and of class C^2, it can be assumed that S is given in the form S: $z = z(x, y)$, where (x, y) varies on a small domain, say the circle $x^2 + y^2 < \epsilon^2$, on which the function $z(x, y)$ is of class C^2. It can also be assumed that the oriented unit normal $N = (\xi, \eta, \zeta)$ of S at $(x, y) = (0, 0)$ is in the positive half of the z-axis, hence $\zeta > 0$ on $x^2 + y^2 < \epsilon^2$ (if ϵ is small enough). Then, if

$$q = \xi/\zeta, \qquad r = \eta/\zeta, \tag{27}$$

it is seen from (26), where $p = X \cdot N$ and $X = (x, y, z)$, that

$$qx + ry = H + L, \tag{28}$$

where

$$H = H(q, r, 1) \text{ and } L = -z = -z(x, y) \equiv L(x, y) \tag{29}$$

are functions of class C^1 and C^2, respectively. But (28) and (29) imply that

$$x = H_q, \qquad y = H_r, \tag{30}$$

which, together with (28), defines a Legendre transformation.

In fact, in view of the involutory character of a Legendre transformation, (30) and (28) together are equivalent to

$$q = L_x, \qquad r = L_y \tag{31}$$

and (28) together, provided that $L = L(x, y)$ is a function of class C^2 and of non-vanishing Hessian (i. e., provided that the transformation (31) of (x, y) into (q, r) is of class C^1 and of non-vanishing Jacobian). But the definition (29) of L shows that this proviso is satisfied, since the function $z(x, y)$ is of class C^2 and has a Hessian which, being a positive multiple of the curvature $K = K(x, y)$, does not vanish by assumption.

It now follows from a known fact on Legendre transformations (cf. [7], p. 6) that the transformation (30) of (q, r) into (x, y) is of class C^1 (and of non-vanishing Jacobian). But this means that $H = H(q, r, 1)$ in (30) is a function of class C^2 in (q, r). It follows therefore from (27) and (26) that $H(\xi, \eta, \zeta)$ is a function of class C^2 in $(\xi, \eta, \zeta) \neq (0, 0, 0)$.

From the formal point of view, the supporting function of a parallel surface results by adding a constant to the supporting function of the given surface. It is clear that, because of this formal circumstance, the fact just verified could be applied in the direction of (iii).

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] S. Bernstein, "Sur la déformation des surfaces," *Mathematische Annalen*, vol. 60 (1905), pp. 434-436.

[2] L. Bianchi, *Lezioni di geometria differenziale*, Piza, 1894.

[3] P. Hartman and A. Wintner, "On the asymptotic curves of a surface," *American Journal of Mathematics*, vol. 73 (1951), pp. 149-172.

[4] L. Lichtenstein, "Ueber den analytischen Charakter der Lösungen regulärer zweidimensionaler Variationsprobleme," *Bulletin International de l'Académie des Sciences de Cracovie, Classe des Sciences Mathématiques et Naturelles*, sér. A, 1913, pp. 915-941.

[5] J. Steiner, *Gesammelte Werke*, vol. 2 (1882), pp. 170-176.

[6] A. Wintner, "On isometric surfaces," *American Journal of Mathematics*, vol. 74 (1952), pp. 192-214.

[7] ———, *The Analytical Foundations of Celestial Mechanics*, Princeton, 1941.

THE SPECTRA OF QUANTUM-MECHANICAL OPERATORS.*

By C. R. PUTNAM.

---

1. Let A and B denote "Hermitian" operators (in a sense to be defined), satisfying the commutator "equation"

$$AB - BA = iI, \tag{1}$$

where, here, and throughout this paper, i is the imaginary unit, I denotes the identity operator, and all operators are to be regarded with respect to an underlying Hilbert space $\mathfrak{H}$. The justification for the appearance of the quotation marks above is that the unqualified relation (1) can *never* hold as an actual equation. Specifically, the identity operator appearing on the right side of (1) is defined for all elements f in $\mathfrak{H}$; thus, if (1) were a true equation, then clearly both A and B would be defined for all f in $\mathfrak{H}$ and hence would be bounded operators,[1] in contradiction to known results; cf. [12], [8], [10]. What is needed, then, is the specification of a domain, say Ω, of elements in $\mathfrak{H}$ for which (1) does hold, that is for which

$$(AB - BA)f = if, \tag{2}$$

for all f in Ω. The statement that "(1) holds on Ω" will then be taken to mean that (2) holds for all f in Ω. It should be noted that Ω belongs both to the domain of definition of the operator AB and to that of BA.

This paper will be devoted to the problem of determining the nature of A and B and their spectra in relation to the domain Ω. Certain results in this direction have recently been obtained by Rellich [8] under assumptions quite different to those of the present paper. In fact, the conditions imposed in [8] are such as to imply that the operators A and B are of the Heisenberg-Schroedinger type, and are therefore both unbounded; whereas in the case at hand, the fact that the boundedness of one of the operators is not precluded is essentially connected with the nature of the theorems to be presented below. For other results bearing on problems somewhat related to those considered in the present paper, see [7].

---

* Received March 3, 1951.

[1] If A is any symmetric operator for which $\mathfrak{D}_A = \mathfrak{H}$ (cf. section 2) then necessarily A is self-adjoint and is bounded ([9], pp. 51, 59).

The significance of the role of Ω in determining the spectral properties of A and B can perhaps be most readily demonstrated if A and B are defined by

$$A = id/dx, \qquad B = x. \tag{3}$$

If f denotes any differentiable function, then clearly (2) holds, so that in a sense, (1) can be regarded as a "formal" identity. First, suppose that $\mathfrak{H}$ consists of all (complex-valued) functions f of class $L^2(-\infty, \infty)$. Then (1) holds on a certain subset Ω of $\mathfrak{H}$, and furthermore it is known ([5], pp. 66 ff.) that A and B can be regarded as self-adjoint operators (as defined in the next section), and that each posssesses the whole line $-\infty < \lambda < \infty$ as its spectrum; in addition, neither A nor B possesses an eigenvalue.

Next, let $\mathfrak{H}$ consist of all functions of class $L^2[a, b]$, where $[a, b]$ denotes a *finite* interval. If the boundary condition $|f(a)/f(b)| = 1$ (among other conditions) is required for the functions in the domain of A, it is known that again both A and B can be regarded as self-adjoint operators. In this case, however, the spectrum of A consists of a discrete sequence of eigenvalues $\lambda_1, \lambda_2, \cdots$, which satisfy $\lambda_n \to \infty$ as $n \to \infty$, while the spectrum of B is the entire interval $a \leq \lambda \leq b$.

It was thought by Dirac ([1], p. 94) that the situation in which the spectra of A and B consist of every real number was a consequence of the "equation" (1). However, as shown by the last example (which has been cited in a quantum-mechanical connection in the physical literature; [2], p. 299) such is, in general, not the case. Moreover, as is shown in Appendix II below, it is quite possible that A and B be formally Hermitian matrices (so that $a_{jk} = \bar{a}_{kj}$ and $b_{jk} = \bar{b}_{kj}$) which satisfy (1), and are such that one of the pair is bounded (and hence possesses a spectrum contained in a finite interval). In fact, the matrix consideration referred to here reflects the situation for the operators (3) on the finite interval.

As previously remarked, the present paper will deal with the problem of characterizing the operators A and B (both of which, except in Theorem I and the Lemma below, are self-adjoint) satisfying (1), in regard to Ω. A different type of problem, concerning the nature of the spectrum of a certain self-adjoint differential equation and the concomitant boundary value, problem, has been considered by Wintner [13].

The discussion of the cases bearing on (3) considered above, will be considered further in section 6.

2. Terminology. The underlying Hilbert space will be denoted by $\mathfrak{H}$;

it is understood that $\mathfrak{H}$ is not the trivial space consisting of the zero element alone. The definitions and terminology will be that of Nagy [3]. Thus, an operator A will be called symmetric if A is linear and densely defined, and if $(Af, g) = (f, Ag)$ for all elements f, g in $\mathfrak{D}_A$, the domain of A; hence $A \subseteqq A^*$. An operator A will be called normal if A is linear, densely defined, closed, and satisfies $AA^* = A^*A$; in particular, A will be called self-adjoint if A is normal and $A = A^*$. It is to be emphasized that the *unqualified* equation $A = B$ entails both relations $A \subseteqq B$ and $A \supseteqq B$.

By $\mathfrak{R}_A$ will be meant the range of the operator A, so that $\mathfrak{R}_A$ consists of all elements g in $\mathfrak{H}$ such that $g = Af$, f in $\mathfrak{D}_A$. Finally, if Ω denotes any set contained in $\mathfrak{D}_A$, then $A(\Omega)$ denotes the image of Ω, that is, the set of all elements g in $\mathfrak{H}$ such that $g = Af$, f in Ω.

All operators considered in this paper are linear.

3. The theorems and some of the proofs will be given in this section and in section 5; the remaining proofs (those of Theorems I, II and III) will be given in sections 7-10.

THEOREM I. *Let A be symmetric and let B be an arbitrary linear operator in $\mathfrak{H}$ satisfying the commutator equation* (1) *on a set Ω of $\mathfrak{H}$. Then A cannot possess a real eigenvalue λ with an eigenfunction belonging to Ω. Moreover, the above statement remains true if A and B are interchanged.*

It is to be noticed that the above theorem does not preclude the existence of eigenfunctions for the operators A or B; in fact, it is quite possible that both A and B are self-adjoint, satisfy (1) on a set Ω, and that the entire spectrum of A consists of (real) eigenvalues (with corresponding eigenfunctions which, by Theorem I, cannot belong to Ω). Such an example is furnished by (3) with reference to the finite interval $[a, b]$; cf. the remarks in section 1. The situation is readily explained in the present instance by the fact that if a function $f(x)$ is in $\mathfrak{D}_A$, it satisfies the boundary condition $|f(a)/f(b)| = 1$. Since, however, $\Omega \subseteqq \mathfrak{D}_{AB}$, all functions of Ω necessarily satisfy this same condition; in particular if f is an eigenfunction of A belonging to Ω, then both $f(x)$ and $xf(x)$ would have to satisfy the boundary condition. This is impossible since $|f(a)| = |f(b)| \neq 0$ for such f's.

Theorem I is similar to a portion of Theorem III of [7].

The next theorem is

THEOREM II.[2] *Let A and B denote self-adjoint operators in $\mathfrak{H}$ satisfying*

---

[2] The author wishes to thank Professor Kodaira for several valuable comments and suggestions in connection with the formulation and proof of this theorem.

(1) *on a set* Ω, *and suppose that the set* $AB(\Omega)$ *is dense in* $\mathfrak{H}$. *In addition, let* A^{-1} *be bounded. Then* (i) $\mathfrak{D}_A \subseteq \mathfrak{R}_B$ *and* (ii) A *is unbounded. Moreover, the theorem remains true if* A *and* B *are everywhere interchanged.*

The next theorem is similar to the above but with a slightly altered hypothesis.

THEOREM III. *The assertions of Theorem* II *above remain valid if the assumption that* $AB(\Omega)$ *be dense in* $\mathfrak{H}$ *is replaced by the assumption that* $\Phi = BA(\Omega)R_{AB}$ *be dense in* $\mathfrak{H}$.

In the next section, a lemma will be proved which will be useful in deriving certain consequences of the above theorems.

4. For an arbitrary pair of complex numbers λ, μ and an arbitrary pair of operators A, B, define A_λ and B_μ by

$$A_\lambda = A - \lambda I, \qquad B_\mu = B - \mu I. \tag{4}$$

It is readily seen that if λ denotes a real number, then the operator A_λ is self-adjoint (symmetric) if and only if A is self-adjoint (symmetric). First, the following Lemma will be derived:

LEMMA. *If* A *and* B *denote arbitrary operators* (*not necessarily self-adjoint or even symmetric*) *in a Hilbert space* $\mathfrak{H}$, *and if* (1) *holds on* Ω, *then*

$$A_\lambda B_\mu - B_\mu A_\lambda = iI \qquad (\lambda \text{ and } \mu, \text{ arbitrary complex numbers}), \tag{5}$$

also holds on Ω.

Proof. Since $\Omega \subseteqq \mathfrak{D}_{AB} \subseteqq \mathfrak{D}_B$, and $\Omega \subseteqq \mathfrak{D}_{BA} \subseteqq \mathfrak{D}_A$, it is clear from (1) that

$$AB - \mu A - \lambda B + \lambda\mu I - BA + \mu A + \lambda B - \lambda\mu I = iI \tag{6}$$

holds on Ω. Since $AB + AC \subseteqq A(B + C)$, and $(A + B)C = AC + BC$ hold for arbitrary operators A, B, C (cf. [3], p. 28), relation (6) implies first that

$$A(B - \mu I) - \lambda(B - \mu I) - (B - \mu I)A + \lambda(B - \mu I) = iI$$

holds on Ω, and then that

$$(A - \lambda I)(B - \mu I) - (B - \mu I)(A - \lambda I) = iI$$

holds on Ω. As this last relation is simply (5), the proof of the Lemma is complete.

5. It is clear that if the role of (1) in Theorems I-III is played by (5), and if the role of the set Ω is played by some set $\Omega_{\lambda\mu}$, then theorems similar to I-III, but pertaining to A_λ and B_μ, where λ and μ are real, may readily be obtained. According to the Lemma, the set $\Omega_{\lambda\mu}$, on which (5) holds, contains the set on which (1) holds, that is, $\Omega \subseteq \Omega_{\lambda\mu}$. Since another application of the Lemma shows that $\Omega_{\lambda\mu} \subseteq (\Omega_{\lambda\mu})_{-\lambda-\mu}\ (=\Omega)$, it follows, however, that even $\Omega = \Omega_{\lambda\mu}$ (λ, μ complex, arbitrary) is valid.

In addition it can be remarked that in order to prove, for instance, that A is unbounded, it is sufficient to prove that A_λ is unbounded for some (real or complex) value λ.

For the sake of application to the problem of locating spectra, the following theorem, which involves A_λ and B_μ, and which is a corollary of Theorem II, will be stated.

THEOREM IV. *Let A and B denote self-adjoint operators in $\mathfrak{H}$ satisfying* (1) *on a set Ω, and let A_λ and B_μ be defined by* (4).

(i) *Consider the set, S, of real values λ for which there is a real value $\mu = \mu(\lambda)$ such that the set $A_\lambda B_\mu(\Omega)$ is dense in $\mathfrak{H}$ and $\mathfrak{D}_{A_\lambda}$ is not contained in $\mathfrak{R}_{B_\lambda}$. Then the set S is contained in the spectrum of the operator A.*

(ii) *The above statement* (i) *remains true if the symbols A and B are interchanged throughout.*

Remark. It should be noted that if a self-adjoint operator A possesses the spectral resolution $A = \int_{-\infty}^{\infty} \alpha dE_\alpha$, then $A_\lambda = \int_{-\infty}^{\infty} (\alpha - \lambda) dE_\alpha$. Furthermore, if λ is not an eigenvalue of A, then the inverse of A_λ, that is A_λ^{-1}, exists and is given by $A_\lambda^{-1} = \int_{-\infty}^{\infty} (\alpha - \lambda)^{-1} dE_\alpha$, the resolvent of A. Since A is a self-adjoint (and therefore normal) operator, it can possess at most a denumerable set of eigenvalues λ (for which values A_λ^{-1}, of course, fails to exist). A point λ is *not* in the spectrum of A if and only if A_λ^{-1} exists and is bounded. In addition, it should be noted that if λ is not an eigenvalue, then the inverse operator A_λ^{-1} (as well as A_λ) as defined above is self-adjoint whenever λ is real.

In connection with the above remark see [3], p. 54.

Proof of Theorem IV. Since an interchange of A and B in (1) is equivalent to the substitution of $-A$ for A in this relation, it is clearly sufficient to prove only part (i) of Theorem IV. By virtue of the Lemma, relation

(5) holds on Ω. Let λ be in S, and suppose λ is not in the spectrum of A. Then A_λ^{-1} exists, is self-adjoint, and is bounded. An application of (i) of Theorem II then implies $\mathfrak{D}_{A_\lambda} \subseteq \mathfrak{R}_{B\mu}$, in contradiction to the hypothesis; hence all points of S are in the spectrum of A. This completes the proof of Theorem IV.

6. Examples. Consider the real axis $-\infty < x < \infty$, and on it the self-adjoint operators A and B defined by (3) (cf. section **1**). It is known ([5], pp. 66 ff.) that both A and B possess the set $-\infty < \lambda < \infty$ as their spectra. A proof of this fact, as far as A is concerned, will be given as an application of Theorem IV. Since A is self-adjoint, the spectrum of A is contained on the real line $-\infty < \lambda < \infty$. It will be shown that for any real λ, the function $\mu = \mu(\lambda) \equiv 0$ suffices in part (i) of Theorem IV. Thus, it will first be shown that the set $A_\lambda B(\Omega)$ is dense for all real λ. (It is to be noted that $B_0 = B$.) This set consists of all functions g such that

$$g = if' - \lambda f, \tag{7}$$

where $f = xF$, and F is in Ω. Multiplication of (7) by $e^{i\lambda x}$ yields

$$ge^{i\lambda x} = i(fe^{i\lambda x})'. \tag{8}$$

Let $k = k(x)$ be in $L^2(-\infty, \infty)$ and satisfy

$$\int_{-\infty}^{\infty} kg\,dx = 0 \text{ for all functions } g \text{ of (7)}. \tag{9}$$

Consider only functions F with continuous first derivatives on $-\infty < x < \infty$ and which, in addition, are such that $f \equiv 0$ outside of a finite interval $[a, b]$. Clearly the functions F, as well as the functions $f = xF$, belong to the class Ω and are dense in $\mathfrak{H}$. From $\int_a^b kg\,dx = 0$, one obtains $i\int_a^b ke^{-i\lambda x}(fe^{i\lambda x})'dx = 0$. Put $K(x) = ike^{-i\lambda x}$; then the last relation implies $\int_a^b fe^{i\lambda x}dK(x) = 0$ for all functions f of the type considered above. This clearly implies that $K = \text{const.}$ on $[a, b]$. Since a and b are arbitrary, $K = \text{const.}$ on $-\infty < x < \infty$. Since K belongs to class $L^2(-\infty, \infty)$, K (and hence k) is identically zero. Thus the set $A_\lambda B(\Omega)$ is dense for *every* λ. The assertions regarding A will follow if it is shown that $\mathfrak{D}_{A_\lambda}$ is not contained in $\mathfrak{R}_B$ for any real λ. However, if a function f is in $\mathfrak{R}_B$, then $f = xg$, where $g \subset \mathfrak{D}_B$, and consequently $f/x \subset L^2(-\infty, \infty)$. But the existence of functions f in $\mathfrak{D}_{A_\lambda}$ for which f/x is not in $L^2(-\infty, \infty)$ is clear; in fact, one need only arrange that $f(0) \neq 0$,

so that (f being continuous), $\int_{-\infty}^{\infty} |f/x|^2 dx = \infty$ holds. Consequently, the spectrum of A contains the entire line $-\infty < \lambda < \infty$, and is therefore just this line.

The situation for the line $-\infty < x < \infty$ is to be contrasted with that for the finite interval $[a, b]$. In the latter case, as remarked in section 1, the boundary condition $|f(a)/f(b)| = 1$; (among other requirements) must be imposed on the functions of $\mathfrak{D}_A$. Since, among other things, all functions of Ω must satisfy the boundary condition $f(a) = f(b) = 0$, one can readily see, by direct verification, that the set $\Gamma = A_\lambda B_\mu(\Omega)$ is, in the present case, never dense (λ, μ real, arbitrary), and that the set $\Phi = B_\mu A_\lambda(\Omega)$ is surely not dense if B_μ^{-1} is bounded. In fact, the function $k = e^{i\lambda x}$ is orthogonal to all elements of Γ; while if B_μ^{-1} is bounded, so that μ does not belong to $[a, b]$, then the function $k = e^{i\lambda x}(x - \mu)^{-1}$ is orthogonal to all elements of Φ.

7. *Proof of Theorem* I. It is sufficient to prove the theorem in the case that A is symmetric and B is arbitrary; cf. the above remark at the beginning of the proof of Theorem IV. According to the Lemma, (5) holds on Ω (λ and μ arbitrary). Hence, if A possesses a real eigenvalue with an eigenfunction f in Ω, then $A_\lambda f = 0$ for some real λ, and if $\mu = 0$, then (5) implies $A_\lambda Bf - BA_\lambda f = if$, so that, since B is linear, $A_\lambda Bf = if \subset \Omega$. In particular, since $\Omega \subseteqq \mathfrak{D}_{BA_\lambda} \subseteqq \mathfrak{D}_{A_\lambda}$, the element $A_\lambda Bf \subset \mathfrak{D}_{A_\lambda}$. Since λ is real, A_λ is symmetric, and hence $(g, A_\lambda h) = (A_\lambda g, h)$ whenever g, h belong to $\mathfrak{D}_{A_\lambda}$. Clearly $Bf \subset \mathfrak{D}_{A_\lambda}$. Let $g = A_\lambda Bf$ and $h = Bf$; then $(A_\lambda Bf, A_\lambda Bf) = (A_\lambda A_\lambda Bf, Bf) = 0$, since $A_\lambda A_\lambda Bf = A_\lambda(if) = 0$. Thus $A_\lambda Bf = 0$ and hence $f = 0$, in contradiction to the requirement that f be an eigenfunction. This completes the proof of Theorem I.

8. *Proof of* (i) *of Theorem* II. As above, it is sufficient to prove the theorem as stated in terms of A and B (cf. the beginning of the proof of Theorem IV above). It follows from (1) that $2\Im(ABf, f) = (f, f)$ so that

$$\| ABf \| \geq \tfrac{1}{2} \| f \|, \qquad f \text{ in } \Omega. \tag{10}$$

In order to prove (i), it is sufficient to show that for any element g in $\mathfrak{H}$, there exists an element f of $\mathfrak{D}_B$ for which $A^{-1}g = Bf$. Since $AB(\Omega)$ is dense in $\mathfrak{H}$, there exists a sequence $f_1, f_2, \cdots$, where f_n are in Ω, such that $ABf_n \to g$ as $n \to \infty$. By virtue of (10), $f_n \to f \subset \mathfrak{H}$, as $n \to \infty$. But A^{-1} is bounded, so that $Bf_n \to A^{-1}g$; since B is closed, f is in $\mathfrak{D}_B$, and $Bf = A^{-1}g$. This proves (i).

9. *Proof of* (ii) *of Theorem* II. It will be shown that the assumption that A is bounded leads to a contradiction. Suppose, if possible, that A (as well as A^{-1}) is bounded. Since $\mathfrak{D}_A = \mathfrak{H}$ and, by (i), $\mathfrak{D}_A \subseteq \mathfrak{R}_B$, it follows that B^{-1} exists and is self-adjoint (cf. [3], p. 35, (c)). In addition, B^{-1} is also bounded ([9], p. 59. See also the above footnote 1). Let f belong to $AB(\Omega)$, so that $f = ABg$, where $g \subset \Omega$. An application of (1) to $g = B^{-1}A^{-1}f$ yields $f - BAB^{-1}A^{-1}f = iB^{-1}A^{-1}f$. Let α and β denote the bounded operators $A^{-1}B^{-1}$ and $B^{-1}A^{-1}$ respectively. Then an application of the operator α to the last equation gives

$$\alpha f - \beta f = i\alpha\beta f, \tag{11}$$

which is valid for all f in $AB(\Omega)$ (and hence for all f in $\mathfrak{H}$). Clearly $\alpha = \beta^*$; therefore if U is defined by

$$U = \beta + iI, \tag{12}$$

relation (11) implies

$$U^*U = I. \tag{13}$$

For any bounded operator C, let $\| C \|$ be defined by $\| C \| = \text{l. u. b.} \| Cf \|$ when $\| f \| = 1$; since $(Uf, Uf) = (f, U^*Uf) = \| f \|^2$ for all $f \subset \mathfrak{H}$, it follows that $\| U \| = 1$ (and hence $\| U^* \| = 1$). Moreover, the spectrum of any bounded operator C is contained in the circle $| z | \leq \| C \|$ of the complex plane (cf. [11], p. 143). Hence, by (12), the spectra of β and β^* lie in circles of radii 1 with centers at $z = -i$ and at $z = i$, respectively.

Since $A^{-1}B^{-1} = A^{-1}(B^{-1}A^{-1})A$, it follows from the fact that A and A^{-1} are bounded that β and β^* $(= \alpha)$ have identical spectra (cf. [12]). Consequently, the origin $z = 0$ is the only possible point in the spectra of β and β^*. (Incidentally, that $z = 0$ actually is in the spectrum of β is clear from the fact that the spectrum of a bounded operator is never empty; cf. [11], pp. 143-144.) Hence by (12), $\lambda = 0$ certainly is not in the spectrum of U, so that U is non-singular (possesses a unique, two-sided, bounded inverse); since, as was seen above, $U^*U = I$, it follows that $UU^* = I$ must also hold. Thus U is unitary and β is normal, so that $\beta = B^{-1}A^{-1}$ is the (bounded) zero operator. (It should be noted that the fact that $z = 0$ is the only point in spectrum of a *non-normal* bounded operator C is *not* sufficient to guarantee that C is the zero operator.) It follows that β possesses $\lambda = 0$ as an eigenvalue with any $f \neq 0$ in $\mathfrak{H}$ as a corresponding eigenfunction. One thereby obtains a contradiction to the fact that β is the inverse of AB (cf. [3], p. 28). In fact, $0 = \beta ABf = f$ for all f in $\mathfrak{D}_{AB}$; the existence of some $f \neq 0$ in $\mathfrak{D}_{AB}$, however, follows from the assumption that $AB(\Omega)$ is dense in the (non-trivial) space $\mathfrak{H}$. This contradiction completes the proof of Theorem II.

10. *Proof of Theorem* III. Since $\mathfrak{R}_B$ is dense, B^{-1} exists and is self-adjoint; [3], p. 35, (c). Let $f \subset \Phi$; then $f = BAg$, $g \subset \Omega$. An application of (1) to $g = A^{-1}B^{-1}f$ yields $ABA^{-1}B^{-1}f - f = iA^{-1}B^{-1}f$. As in section 9, let $\alpha = A^{-1}B^{-1}$ and $\beta = A^{-1}B^{-1}$ (with the understanding that α and β need not be bounded in the present case). Since f is in Φ, both elements f and $ABA^{-1}B^{-1}f$ belong to $\mathfrak{D}_\beta$, so that $A^{-1}B^{-1}f$ must also belong to $\mathfrak{D}_\beta$. An application of β to the above equation gives

$$\alpha f - \beta f = i\beta\alpha f, \qquad f \text{ in } \Phi.$$

Since Φ is dense, both α and β are densely defined and, consequently, the adjoints α^* and β^* exist. Since A^{-1} is bounded, $\alpha^* = \beta$ ([3], top of p. 29). Hence β, being an adjoint, is closed. In addition, $\alpha \subseteqq \beta^*$; [3], top of p. 29. The last formula line now implies $\beta^* f - \beta f = i\beta\beta^* f$, for all f in Φ. If, as in section 9, U is defined by (12), it is clear that U is linear, closed, and densely defined, that U^* exists, and that (cf. [3], top of p. 29) $U^* = \beta^* - iI$. It follows that

$$UU^*f = f, \qquad f \text{ in } \Phi. \tag{14}$$

Furthermore $U^{**} = U$ ([3], middle of p. 29). Since (14) holds for all f in Φ, $\Phi \subseteqq \mathfrak{D}_{U^{**}U^*} \subseteqq \mathfrak{D}_{U^*}$. Let h be an arbitrary element in $\mathfrak{H}$; since Φ is dense, there exists a sequence of elements f_n in Φ such that $f_n \to h$ as $n \to \infty$. Clearly

$$(U^*(f_n - f_m), U^*(f_n - f_m)) = ((f_n - f_m), U^{**}U^*(f_n - f_m)),$$

so that by (14) and the equation $U^{**} = U$, $U^*(f_n - f_m) \to 0$. Thus $U^*f_n \to k$, where k is some element of $\mathfrak{H}$. But U^*, being an adjoint, is closed; hence, $U^*h = k$, that is, U^*h is defined for all h in $\mathfrak{H}$. Furthermore, it is readily seen that $(U^*f, U^*f) = (f, U^{**}U^*f) = (f, f)$ for f in Φ. Thus $|U^*f| = |f|$ whenever f is in Φ, and hence (by the above results) whenever f is in $\mathfrak{H}$. Consequently, U^* is bounded; this implies that $U (= U^{**})$ is bounded. It now follows from (12) that β is bounded, and therefore $\mathfrak{R}_{A^{-1}} \subseteqq \mathfrak{D}_{B^{-1}}$, that is, $\mathfrak{D}_A \subseteqq \mathfrak{R}_B$. This proves the first part of Theorem II. The proof of the second part, namely that A is unbounded, is similar to that given in the latter part of section 9. Thus, the assumption that A is bounded readily implies that $U^*U = I$ as well as $UU^* = I$ (the last relation having been obtained above), so that U is unitary. The existence of an $f \neq 0$ in $\mathfrak{D}_{AB}$ follows from the fact that $\Phi \subseteqq \mathfrak{R}_{AB}$, and one obtains a contradiction as before. This completes the proof of Theorem III.

Appendix.

I. Certain formulations of the statement "(1) holds on Ω" can even specify explicitly the set Ω. For instance, if the assumption that (1) holds on Ω is replaced by $AB \subseteqq BA + iI$, then $\Omega = \mathfrak{D}_{AB}$. Furthermore, if it is assumed that $AB = BA + iI$, then $\Omega = \mathfrak{D}_{AB} = \mathfrak{D}_{BA}$. Thus under such revised assumptions, the formulations of the various theorems of this paper can be stated more simply; these revised forms are clear, however, and will not be stated explicitly.

Also, slight variations of the above theorems and even sharpened statements can be obtained in certain instances. Such a case in point is offered by part (ii) of Theorem II. In fact, according to Theorem III of [7], if A is arbitrary and B is normal (not necessarily self-adjoint), then the relation $AB \subseteqq BA + iI$ implies, without further assumptions, that A is unbounded.

II. The existence of a pair of matrices $A = (a_{mn}), B = (b_{mn})$ satisfying relation (1), the Hermitian requirements

$$(a_{mn}) = (\bar{a}_{nm}), \qquad (b_{mn}) = (\bar{b}_{nm}), \tag{15}$$

and

$$\sum_m |a_{mn}|^2 < \infty, \qquad \sum_m |b_{mn}|^2 < \infty, \qquad n = 1, 2, \cdots, \tag{16}$$

and such that one of the pair is bounded, will be shown. According to (16), the rows (hence, by (15), the columns) of A and B belong to Hilbert space; as a consequence, the products AB and BA occurring in (1) (the latter relation being now regarded as an ordinary matrix equation) surely exist. Specifically it will be shown that the pair A, B defined by

$$a_{mn} = \begin{cases} 4imn/\pi(m^2 - n^2), & m \text{ odd (even) and } n \text{ even (odd)} \\ 0, & \text{otherwise} \end{cases}$$

$$b_{mn} = \begin{cases} \pi/2, & m = n \\ -8mn/\pi(m^2 - n^2)^2, & m \text{ odd (even) and } n \text{ even (odd)} \\ 0, & \text{otherwise} \end{cases} \tag{17}$$

satisfies (1), (15) and (16), and that B is bounded. In principle, the verification of these claims can be made directly; however, the proof, which will be given below, will depend on a correspondence between matrices and operators.

Consider the operators $A = id/dx$ and $B = x$, which satisfy the "formal" identity

$$AB - BA = i, \tag{18}$$

as operating on the class, $\mathfrak{F}$, of functions $f(x)$, which are absolutely continuous on the *finite* interval $a \leqq x \leqq b$ and satisfy

$$(19) \qquad f(a) = f(b) = 0, \qquad \int_a^b |f'|^2 dx < \infty \qquad (' = d/dx).$$

It is clear that the set $\mathfrak{F}$ is dense in the Hilbert space $\mathfrak{H}$ of all functions f on $[a, b]$ for which $\int_a^b |f|^2 dx < \infty$, and that, as is seen from (18) and (19),

$$\int_a^b f(\overline{Cg})\,dx = \int_a^b (Cf)\bar{g}\,dx, \text{ where } C = A \text{ or } B \text{ and } f, g \text{ are in } \mathfrak{F}$$

(cf. [5], pp. 66, 68, and [8], p. 111). If desired, the domain of definition of the (unbounded) operator A can be extended so that A becomes a closed operator (cf. [5], pp. 76, 79); it should be noted that B is a bounded operator (and is defined for all f in $\mathfrak{H}$). If $\phi_1(x), \phi_2(x), \cdots$ denotes any orthonormal complete sequence of functions of class $\mathfrak{F}$, and if the matrices $A = (a_{mn})$ and $B = (b_{mn})$ are defined by

$$a_{mn} = i\int_a^b \phi_m \bar{\phi}_n'\,dx, \qquad b_{mn} = \int_a^b x\phi_m\bar{\phi}_n dx,$$

then A and B satisfy (15) and (16), and B is bounded (cf. [11], pp. 184 ff., and [4], pp. 122 ff.). In addition, as a consequence of (18) and (19), it is seen by direct verification, that (1) also holds. By choosing $[a, b]$ to be $[0, \pi]$ and $\phi_n(x) = (2/\pi)^{\frac{1}{2}} \sin nx$, one obtains the matrices A and B of (17). Incidentally, it can be remarked that the spectrum of the matrix B defined by (17) is exactly the interval $[0, \pi]$ ([11], pp. 184 ff.).

By choosing different sequences $\phi_1(x), \phi_2(x), \cdots$ on $[a, b]$, one can obtain various pairs A, B satisfying (1), (15) and (16), and such that B is bounded. All of the bounded matrices B obtained in this manner, are unitarily equivalent to each other ([11], pp. 184 ff. and [4], pp. 122 ff.) and none, of course, is unitarily equivalent to either of the linear oscillator matrices occurring in physics (cf., e. g., [2], p. 370). It will remain undecided whether *all* bounded matrices B, which have identical spectra and which can occur in pairs A, B satisfying (1), (15) and (16), are unitarily equivalent to each other, or what amounts to the same thing, to a fixed such matrix, say the one defined by the second relation of (17), in case the spectrum is the interval $[0, \pi]$.

THE INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[1] P. A. M. Dirac, *The Principles of Quantum Mechanics*, Oxford, 1935 (2nd edition).

[2] E. H. Kemble, *The Fundamental Principles of Quantum Mechanics*, New York and London, 1937.

[3] B. v. Sz. Nagy, *Spektraldarstellung linearer Transformationen des Hilbertschen Raumes*, Ann Arbor, 1947.

[4] J. v. Neumann, "Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren," *Mathematische Annalen*, vol. 102 (1930), pp. 49-131.

[5] ———, *Mathematische Grundlagen der Quantenmechanik*, Berlin, 1932.

[6] C. R. Putnam, "On commutators of bounded matrices," *American Journal of Mathematics*, vol. 73 (1951), pp. 127-131.

[7] ———, "On normal operators in Hilbert space," *ibid.*, vol. 73 (1951), pp. 357-362.

[8] F. Rellich, "Der Eindeutigkeitssatz für die Lösungen der quantenmechanische Vertauschungsrelationen," *Göttinger Nachrichten*, 1946, pp. 107-115.

[9] M. H. Stone, *Linear Transformations in Hilbert Space*, New York, 1932.

[10] H. Wielandt, "Ueber die Unbeschränkheit der Operatoren der Quantenmechanik," *Mathematische Annalen*, vol. 121 (1949), p. 21.

[11] A. Wintner, *Spektraltheorie der unendlichen Matizen*, Leipzig, 1929.

[12] ———, "The unboundedness of quantum-mechanical matrices," *Physical Review*, vol. 71 (1947), pp. 738-739.

[13] ———, "On Dirac's theory of continuous spectra," *ibid.*, vol. 73 (1948), pp. 781-785.

ON NON-OSCILLATORY LINEAR DIFFERENTIAL EQUATIONS OF SECOND ORDER.*

By PHILIP HARTMAN.**

---

1. *Necessary conditions.* In the differential equation

$$y'' + q(t)y = 0, \tag{1}$$

let $q = q(t)$ be real-valued and continuous on $0 \leq t < \infty$. Only real solutions $y = y(t) \not\equiv 0$ will be considered below. The equation (1) will be called non-oscillatory if every solution has only a finite number of zeros on $0 \leq t < \infty$.

(I) *If* $q(t)$ *satisfies*

$$q(t) \to 0, \text{ as } t \to \infty, \tag{2}$$

or, more generally, if

$$\underset{0<v<\infty}{\text{l. u. b.}} \left| \int_u^{u+v} q(t)\,dt \right| / (1+v) \to 0, \text{ as } u \to \infty, \tag{3}$$

then a necessary condition that (1) *be non-oscillatory is that either*

$$\int_0^T q(t)\,dt \to -\infty, \text{ as } T \to \infty, \tag{4}$$

or

$$\int_0^\infty q(t)\,dt = \lim_{T\to\infty} \int_0^T q(t)\,dt \text{ exists} \tag{5}$$

(*as a finite limit*).

In the simple case where the assumption (2) or (3) is replaced by $q \geq 0$, it is known that (5) is a necessary condition (Wintner [9], p. 97). In fact, without any restriction on $q(t)$, it is known that

$$\liminf_{T\to\infty} T^{-1} \int_0^T \left(\int_0^t q(s)\,ds \right) dt < \infty$$

---

* Received May 21, 1951.

** John Simon Guggenheim Memorial Foundation Fellow, on leave of absence from The Johns Hopkins University.

is a necessary condition (Wintner [10]). The following theorem contains this result (but not that of (I)):

(II) *A necessary condition that* (1) *be non-oscillatory is that either*

$$\liminf_{T\to\infty} T^{-1}\int_0^T\Big(\int_0^t q(s)\,ds\Big)\,dt = -\infty \tag{6}$$

or

$$\lim_{T\to\infty} T^{-1}\int_0^T\Big(\int_0^t q(s)\,ds\Big)\,dt = M \text{ exists} \tag{7}$$

(*as a finite limit*).

The proof of (II) will show that, in the case (7), the mean value M exists "absolutely" and, in fact,

$$\lim_{T\to\infty} T^{-1}\int_0^T |\,Q(t)|^2 dt = 0, \tag{8}$$

where

$$Q(t) = M - \int_0^t q(s)\,ds. \tag{9}$$

Using the methods of [3], where however it is assumed that (5) holds and (9) reduces to $Q(t) = \int_t^\infty q(s)\,ds$, it can be shown that in the case (7) the function (9) must satisfy

$$\int_0^\infty \exp\Big(-A\int_0^t Q(s)\,ds\Big)\,dt = \infty \tag{10}$$

for $0 < A \leq 4$; but, according to Wintner ([11], § 7), not necessarily for $A > 4$. (That (10) must hold for $A = 2$ when (5) is assumed was first proved by Wintner [11], § 7.)

(II bis) *The case* (6) *of* (II) *is compatible with each of the possibilities*:

$$\limsup_{T\to\infty} T^{-1}\int_0^T\Big(\int_0^t q(s)\,ds\Big)\,dt \text{ is } -\infty, \text{ finite or } +\infty. \tag{11}$$

Of course, (I) rules out the last two possibilities when (2) or (3) holds. These possibilities are also ruled out by somewhat milder assumptions.

(III) *If $q(t)$ is half-bounded on $0 \leq t < \infty$, or more generally, if there exists a positive constant ϵ such that*

$$\int_u^{u+v} q(t)\,dt \text{ is half-bounded for } 0 \leq u < \infty, 0 \leq v \leq \epsilon, \tag{12}$$

then a necessary condition that (1) *be non-oscillatory is that either* (4) *or* (7) *hold.*

2. *Auxiliary theorems.* The proof of (II) will depend on the following:

(A) *If* (1) *is non-oscillatory, then a necessary and sufficient condition that*

$$\int^{\infty} (y'/y)^2 dt < \infty \tag{13}$$

hold for one (and/or every) solution $y = y(t)$ of (1) *is that* (7) *hold.*

This implies the result of Wintner ([11], § 6) that a sufficient condition that (13) hold for every solution of (1) is that (5) be valid. In view of (II), it is seen that even

$$\liminf_{T\to\infty} T^{-1} \int_0^T \left(\int_0^t q(s)\,ds \right) dt > -\infty \tag{14}$$

is sufficient (since (14) implies (7), when (1) is non-oscillatory).

The proof of (I) will depend on the following:

(B) *If* (1) *is non-oscillatory, then* (3) *is necessary and sufficient in order that*

$$y'(t)/y(t) \to 0, \text{ as } t \to \infty, \tag{15}$$

hold for one (and/or every) solution $y = y(t)$ of (1).

This implies the result of Wintner ([11], § 6) that a sufficient condition that (15) hold for every solution of (1) is that (5) be valid.

The theorem (B) is an analogue of a theorem in [2], § 11, where $q(t)$ in (1) is replaced by $q(t) - 1$, and (15) by $y'/y \to \pm 1$, as $t \to \infty$ (but where it is asserted, not assumed, in one half of the theorem that (1) is non-oscillatory.)

Theorem (B) implies the result of Perron [6] that (15) holds for every solution of (1) if (2) holds and $q(t) \leq 0$. Actually, Perron treats the more general equation

$$y'' + p(t)y' + q(t)y = 0, \qquad (p, q \text{ real}) \tag{16}$$

and his assumptions are that (2) holds, $q(t) \leq 0$,

$$p(t) \to 0, \text{ as } t \to \infty, \tag{17}$$

and $p(t) \leq 0$. Since there is no loss of generality in assuming that $p(t)$ is locally smooth, it follows from (B) and the variation of constants

$$y = z \exp(-\tfrac{1}{2} \int_0^t p(s)\,ds),$$

which gives the differential equation $z'' + (q - p^2/4 - p'/2)z = 0$ and the relation $y'/y = z'/z - p/2$, that if (17) is assumed, then (B) remains true if "(16)" is read in place of "(1)".

Since (16) is non-oscillatory when $q \leq 0$, it follows that Perron's assumption $p \leq 0$ is superfluous. This fact has already been noted by Hamel [1] who, under the assumptions (2) and (17), gave necessary and sufficient conditions that (16) be non-oscillatory and that its solutions satisfy (15). But it is clear from the extension of (B) to the differential equation (16) that Hamel's conditions must then be merely necessary and sufficient conditions that (16) be non-oscillatory. From this point of view, Hamel's criterion, based on a variational principle, that (16) be non-oscillatory has nothing to do with his assumptions (2) and (17). This is clear from a glance at the first part of his proof. His criterion has been generalized by Putnam [7].

3. *Proof of* (B). If $y = y(t)$ is any solution of (1), then $y(t)$ does not vanish for large t and its logarithmic derivative

$$x = y'/y \tag{18}$$

satisfies the Riccati differential equation

$$x' + x^2 + q(t) = 0 \tag{19}$$

(for those t for which (18) is defined). Hence, (B) is a consequence of the case $r(x) = -x^2$ of the following lemma, which is an analogue of the theorem (II*) in [2], § 9.

LEMMA 1. *Let $r(x)$ be defined and continuous for $-\infty < x < \infty$ and have the properties that*

$$r(0) = 0, \tag{20}$$

and that there exists, on $0 \leq x < \infty$, *a monotone (non-decreasing) function* $R(x)$ *satisfying*

$$|r(x)| \geq R(|x|) > 0 \text{ for } 0 < |x| < \infty \tag{21}$$

and

$$\int^{\infty} dx/R(x) < \infty. \tag{22}$$

Let $x = x(t)$ *be defined for large* t *and be a solution of*

$$x' = r(x) - q(t). \tag{23}$$

Then (3) *is necessary and sufficient in order that*

$$x(t) \to 0, \text{ as } t \to \infty. \tag{24}$$

Proof. The necessity of (3) follows from the inequality

$$\left|\int_u^{u+v} q(t)\,dt\right| \leq |x(u+v) - x(u)| + \int_u^{u+v} |r(x(t))|\,dt,$$

where (20), (21) and (24) imply that $r(x(t)) \to 0$, as $t \to \infty$.

In order to prove the sufficiency of (3), it will first be shown that (3) implies

$$\liminf_{t\to\infty} x(t) \leq 0 \leq \limsup_{t\to\infty} x(t). \tag{25}$$

Suppose, if possible, that the second inequality does not hold; so that

$$x(t) \leq -c < 0 \tag{26}$$

for large t, say for $a \leq t < \infty$. Then, by (21) and (23), for $a \leq u < +v < \infty$,

$$|x(u+v) - x(u)| \geq R(c)v - \left|\int_u^{u+v} q(t)\,dt\right|. \tag{27}$$

Clearly, (3) implies that, if u is sufficiently large and fixed, then the right side of (27) tends to ∞, as $v \to \infty$. Hence (26) shows that $x(t) \to -\infty$, as $t \to \infty$. From (23),

$$x(t) = \int_a^t r(x(s))\,ds - \int_a^t q(s)\,ds + x(a).$$

Since (21) and (22) imply that $|r(x(t))| \to \infty$, as $t \to \infty$, the condition (3) and the last formula line imply $r(x(t)) \to -\infty$, as $t \to \infty$. Hence (3) gives

$$-x(t) \geq \tfrac{1}{2} \int_a^t R(-x(s))\,ds \text{ for large } t.$$

Since R is monotone, it follows that

$$2S' \geq R(S), \text{ where } S(t) = \tfrac{1}{2} \int_a^t R(-x(s))\,ds$$

tend to ∞, as $t \to \infty$. A quadrature gives

$$2 \int^{S(t)} dS/R(S) \geq t + \text{const.} \tag{28}$$

If $t \to \infty$, this relation contradicts (22). This proves the second inequality in (25); the first is proved similarly.

Suppose, if possible, that the limit superior, as $t \to \infty$, of $x(t)$ is positive, say has the value $3c > 0$. Then, by (25), there exist large values of u and $u + v (> u)$ such that $c \leq x(t) \leq 2c$ for $u \leq t \leq u$ and that $x(u+v) - x(u) = \pm c$, where $\pm$ is $-\operatorname{sgn} r(c)$. Then an analogue of (27) gives

$$\left| \int_u^{u+v} q(t)\,dt \right| \geq c + R(c)v \geq c_0(1+v),$$

where $c_0 = \min(c, R(c)) > 0$. But this contradicts (3). Hence, the limit superior of $x(t)$ is 0; similarly, it is proved that the limit inferior is 0. This completes the proof of (24) and of Lemma 1.

4. *Proof of* (A). If $y = y(t)$ is any solution of (1), then its logarithmic derivative (18) satisfies the Riccati equation (19) for large t, say for $a \leq t < \infty$. An integration of (19) gives

$$x(t) + \int_a^t x^2(s)\,ds = x(a) - \int_a^t q(s)\,ds \tag{29}$$

for $t \geq a$.

Suppose that (13) holds, it will be shown that (7) holds. The assumption (13) implies that (29) can be writen in the form

$$x(t) - \int_t^\infty x^2(s)\,ds = M - \int_0^t q(s)\,ds, \tag{30}$$

where

$$M = x(a) - \int_a^\infty x^2(s)\,dc + \int_0^a q(s)\,ds. \tag{31}$$

Since (13) implies

$$T^{-1}\int_a^t x^2(t)\,dt \to 0, \text{ as } T \to \infty,$$

and

$$T^{-1}\int_a^T \Big(\int_t^\infty x^2(s)\,ds\Big)^2 dt \to 0, \text{ as } T \to \infty,$$

the relation (8) follows from (9) and (30). In particular, (7) holds (Hölder), as was to be proved.

In order to prove the converse, note that (29) and a quadrature give

$$t^{-1}\int_a^t x(s)\,ds + t^{-1}\int_a^t \Big(\int_a^s x^2(u)\,du\Big)ds = t^{-1}x(a)\,(t-a) - t^{-1}\int_a^t \Big(\int_a^s q(u)\,du\Big)ds. \tag{32}$$

The assumption (7), or even (14), means that the right side is bounded from above, as $t \to \infty$. If (13) fails to hold, then

$$-t^{-1}\int_a^t x(s)\,ds \geq \tfrac{1}{2}t^{-1}\int_a^t \Big(\int_a^s x^2(u)\,du\Big)ds \text{ for large } t.$$

Then Schwarz's inequality

$$|\,t^{-1}\int_a^t x(s)\,ds\,| \leq (t^{-1}\int_a^t x^2(s)\,ds)^{\frac{1}{2}}$$

shows that

$$4t\int_a^t x^2(u)\,du \geqq \Big(\int_a^t \Big(\int_a^s x^2(u)\,du\Big)ds\Big)^2 \text{ for large } t.$$

This is equivalent to

$$4tS' \geq S^2, \text{ where } S(t) = \int_a^t \Big(\int_a^s x^2(u)\,du\Big)ds$$

tends to ∞, as $t \to \infty$. A quadrature of this inequality gives

$$4 \int^{S(t)} ds/S^2 \geq \log t + \text{const.}$$

Since this is impossible for large t, it follows that (7), or even (14), implies (13) for every solution $y(t)$ of (1). This completes the proof of (A).

5. *Proof of* (II). Suppose that (1) is non-oscillatory and that (6) does not hold (that is, that (14) holds). It must be shown that the alternative, (7), holds. By the proof just completed, (14) implies (13) for every solution $y(t)$ of (1). But (13), for one solution, implies (7), by (A). This proves (II).

6. *Proof of* (III). If (4) fails to hold, then the integral in (4) exceeds some fixed constant when $t = t_n$, where $t_1 < t_2 < \cdots$ is some unbounded sequence of t-values. The half-boundedness of q, or the more general condition (12), implies the existence of constants C and $\epsilon > 0$ such that

$$\int_0^t q(s)\,ds > -C \tag{33}$$

for $t_n < t < t_n + \epsilon$ or $t_n - \epsilon < t < t_n$, where $n = 1, 2, \cdots$, and C, ϵ are independent of n. The assertion (III) is contained, therefore, in the following:

LEMMA 2. *If* (1) *is non-oscillatory and if there exists a constant* C *such that* (33) *holds for all* t *on a set* E *of infinite measure, then* (7) *is valid.*

Proof. By virtue of (A), it is sufficient to show that (13) holds for a solution $y = y(t)$ of (1). The Ricatti equation (19) and its integrated form (29) are satisfied for $x = y'/y$ for large t, say $a \leq t < \infty$. By (33), the right side of (29) is bounded from above if t is on the set E. Hence, if (13) fails to hold, a modification of the arguments between (27) and (28) leads to the contradiction

$$2 \int_{E(T)} dS(t)/S^2(t) \geq \text{measure of } E(T),$$

where T is large, $E(T)$ denotes the portion of E on $T \leq t < \infty$ and

$$S(t) = \tfrac{1}{2} \int_a^t x^2(s)\,ds \to \infty, \text{ as } t \to \infty.$$

This proves Lemma 2.

7. *Proof of* (I). If (1) is non-oscillatory and (2), or more generally (3), hold then, either (4) or (7) holds, by (III). It remains to show that when (7) holds, (5) is valid.

Suppose that (7) holds, then (13) is true for every solution of (1), by (A). In addition, (15) is true, by (B). The equation (30), which is valid for large t, then implies (5). This proves (I).

8. *Other necessary conditions.* Changes of the variables in (1), followed by applications of the above theorems lead to different necessary conditions that (1) be non-oscillatory. For example, if the new independent variable, $s = t^\gamma$, where $\gamma > 0$, and the new dependent variable $z = t^{\frac{1}{2}(\gamma-1)}y$ are introduced, then (1) becomes

$$d^2z/ds^2 + \gamma^{-2}[q(t)t^{2-2\gamma} - (1-\gamma^2)t^{-2\gamma}/4]z = 0, \text{ where } t = s^{1/\gamma}.$$

An application of (II) shows that if

$$\liminf_{T\to\infty} \int^T q(t)t^{1-\gamma}dt > -\infty,$$

then a necessary condition that (1) be non-oscillatory is that the mean value

$$\lim_{T\to\infty} T^{-\gamma} \int^T t^{\gamma-1}\Big(\int^t q(u)u^{1-\gamma}du\Big)dt = M \text{ exist} \tag{34}$$

(absolutely). In particular, if $q \geq 0$, then

$$\int^\infty q(t)t^{1-\gamma}dt < \infty \text{ for all } \gamma > 0.$$

This result can be deduced from Wintner [9], p. 97; [10]; [11]; Hille [4], p. 243; and Leighton [5]. If $\gamma > 0$ is fixed and $q(t)t^{2-2\gamma} \to 0$, as $t \to \infty$, then (I) show that $\lim \int^T q(t)t^{1-\gamma}dt$, as $T \to \infty$, exists either as a finite number or as $-\infty$. Finally, if $\gamma \geq 1$ and $q(t)t^{1-\gamma}$ is half-bounded, as $t \to \infty$, then it can be shown, by (III), that either (34) holds or $\int^T q(t)t^{1-\gamma}dt \to -\infty$, when (1) is non-oscillatory.

9. *Proof of* (II bis). The first possibility in (11) is realized, for example, by $q(t) \equiv -1$.

In order to obtain examples illustrating the other two possibilities of (11), suitable continuously differentiable functions $x = x(t)$ will be constructed on $0 \leq t < \infty$. The function $q(t)$ will then be defined, in terms of $x(t)$, by the Riccati equation (19). It then follows that the corresponding differential equation (1) is non-oscillatory; in fact, $y(t) = \exp \int_0^t x(s)\,ds$ is a non-vanishing solution of (1) for $0 \leq t < \infty$. This type of construction of counter-examples in the theory of the differential equation (1) has been given in [8], p. 394 and [2], § 21.

The function $x = x(t)$ to be defined now will be smooth except for a sequence of discontinuities, which tend to ∞. It will be clear from the proof that these discontinuities can easily be avoided.

Since (32) follows from (19), where a is arbitrary, it must be shown that if $X(t)$ denotes the left side of (32), then for suitable choices of $x = x(t)$,

$$\limsup_{t\to\infty} X(t) = \infty \tag{35}$$

and

$$\liminf_{t\to\infty} X(t) \text{ is finite or } -\infty. \tag{36}$$

The relation (35) is equivalent to (6); while the first and second alternatives of (36) are equivalent to the second and third, respectively, of (11).

Let $\epsilon_1, \epsilon_2, \cdots$ be a sequence of positive numbers satisfying

$$\alpha > \epsilon_n \to 0, \text{ as } n \to \infty, \qquad (\alpha = 1/e). \tag{37}$$

Define $x(t)$, on $(a =)\, 1 - \alpha \leq t < \infty$, by placing $x(t) = 1/(n-t)\log(n-t)$ or $x(t) = 0$ according as $n - \alpha \leq t < n - \epsilon_n$ or $n - \epsilon_n \leq t < n + 1 - \alpha$ for $n = 1, 2, \cdots$. Then

$$T^{-1} \int_a^T x(t)\,dt = -T^{-1} \sum_{k=1}^{n} \log\log \epsilon_k^{-1} \text{ for } n - \epsilon_n \leq T \leq n + 1 - \alpha. \tag{38}$$

The identity

$$\int_a^T \Big(\int_a^t x^2(s)\,ds \Big)\,dt = \int_a^T (T - s)x^2(s)\,ds$$

and a simple calculation show that, for $n - \epsilon_n \leq T \leq n + 1 - \alpha$,

$$(39) \quad T^{-1} \int_a^T \Big(\int_a^t x^2(s)\,ds \Big) dt = \sum_{k=1}^{n} (1 - k/T) I(\epsilon_k) + T^{-1} \sum_{k=1}^{n} (1 + \log^{-1} \epsilon_k),$$

where

$$(40) \quad I(\epsilon) = \int_\epsilon^a dt/t^2 \log^2 t \sim 1/\epsilon \log^2 \epsilon, \text{ as } \epsilon \to 0.$$

If $T = n + 1 - \alpha$, then $1 - k/T \geq (1 - a)/T$ for $k = 1, \cdots, n$. Also, the second term on the right of (39) is positive, by (37). Hence, $X(T)$, which is the sum of (38) and (39), exceeds

$$T^{-1} \sum_{k=1}^{n} (1 - \alpha) I(\epsilon_k) - \log \log \epsilon_n^{-1}$$

for $T = n + 1 - \alpha$. In view of (40), this implies

$$\text{l. u. b.}_{(n-\alpha, n+1-\alpha)} X(t) \to \infty, \text{ as } n \to \infty;$$

and so, (35) holds.

Let $T = n$, then the contribution of the interval $n - \alpha \leq t \leq n - \epsilon_n$ to $X(T)$ is $n^{-1}(1 + \log^{-1} \epsilon_n - \log \log \epsilon_n^{-1})$. This expression tends to 0 or to $-\infty$ according as ϵ_n tends to 0 sufficiently slowly or sufficiently rapidly, as $n \to \infty$. The contribution of $a \leq t < n - \alpha$ to $X(T)$, where $n - \alpha \leq t \leq n + 1 - \alpha$, depends only on $\epsilon_1, \cdots, \epsilon_{n-1}$ (and not on ϵ_n) and exceeds

$$T^{-1} \sum_{k=1}^{n-1} [(1 - \alpha) I(\epsilon_k) - \log \log \epsilon_k^{-1}] \to \infty, \text{ as } n \to \infty.$$

Consequently, if $c_1, c_2, \cdots$ is any given sequence of numbers, which is bounded from above, then there exists a suitable sequence $\epsilon_1, \epsilon_2, \cdots$ satisfying (37) and

$$\text{g. l. b.}_{(n-\alpha, n+1-\alpha)} X(t) \sim c_n, \text{ as } n \to \infty;$$

in particular,

$$\liminf_{t \to \infty} X(t) = \liminf_{n \to \infty} c_n.$$

On choosing c_n independent of n or $c_n = -n$, the alternatives of (36) are realized. This completes the proof of (II bis).

PARIS, FRANCE.

REFERENCES.

[1] G. Hamel, "Ueber das infinitäre Verhalten der Integrale einer linearen Differentialgleichung zweiter Ordnung, wenn die characterische Gleichung zwei gleiche Wurzeln hat," *Mathematische Zeitschrift*, vol. 1(1918), pp. 220-228.

[2] P. Hartman, "Unrestricted solution fields of almost-separable differential equations," *Transactions of the American Mathematical Society*, vol. 63 (1948), pp. 560-580.

[3] ———, "On linear second order differential equations with small coefficients," *American Journal of Mathematics*, vol. 73 (1951), pp. 955-962.

[4] E. Hille, "Non-oscillation theorems," *Transactions of the American Mathematical Society*, vol. 64 (1948), pp. 234-252.

[5] W. Leighton, "The detection of the oscillation of solutions of a second order linear differential equation," *Duke Mathematical Journal*, vol. 17 (1950), pp. 57-62.

[6] O. Perron, "Ueber das infinitäre Verhalten der Integrale einer linearen Differentialgleichung zweiter Ordnung, wenn die characteristische Gleichung zwei gleiche Wurzeln hat," *Sitzungsberichte der Heidelberger Akademie*, 1917, Abhandlung 9.

[7] C. R. Putnam, "An oscillation criterion involving a minimum principle," *Duke Mathematical Journal*, vol. 16 (1949), pp. 633-636.

[8] A. Wintner, "The adiabatic linear oscillator," *American Journal of Mathematics*, vol. 68 (1946), pp. 385-397.

[9] ———, "On the Laplace-Fourier transcendents occurring in mathematical physics," *ibid.*, vol. 69 (1947), pp. 87-97.

[10] ———, "A criterion of oscillatory stability," *Quarterly Journal of Applied Mathematics*, vol. 7 (1949), pp. 115-117.

[11] ———, "On the non-existence of conjugate points," *American Journal of Mathematics*, vol. 73 (1951), pp. 368-380.

ON THE DIRICHLET SERIES ASSOCIATED WITH RAMANUJAN'S τ-FUNCTION.*

By GEORGE SHAPIRO.[1]

I. Introduction. The Ramanujan function $\tau(n)$ is defined by the identity

$$x\prod_{n=1}^{\infty}(1-x^n)^{24}=\sum_{n=1}^{\infty}\tau(n)x^n;$$

$t(n)$, by the relation

$$t(n)=\tau(n)n^{-11/2}.$$

It is known [7], that

$$t(n)=O(n^{3/10}), \tag{1.1}$$

while Ramanujan has conjectured that

$$t(n)=O(n^{\epsilon}) \text{ for every } \epsilon>0. \tag{1.2}$$

The function $t(n)$ is multiplicative; i. e.,

$$t(mn)=t(m)t(n) \text{ if } (m,n)=1. \tag{1.3}$$

Also,

$$t(p^{\lambda})=t(p)t(p^{\lambda-1})-t(p^{\lambda-2}), \qquad \lambda=2,3,\cdots. \tag{1.4}$$

Put

$$Z(s)=\sum_{n=1}^{\infty}t(n)n^{-s}. \tag{1.5}$$

This series is known to converge for $\sigma>2/5$ and to have as abscissa of absolute convergence $\sigma=1$. Therefore, it follows from (1.3), (1.4) and (1.5) that the Euler representation of $Z(s)$ takes the special form

$$Z(s)=\prod_{p}(1-t(p)p^{-s}+p^{-2s})^{-1} \text{ for } \sigma>1, \tag{1.6}$$

and that the product is absolutely convergent there.

Wilton [9] has shown that

$$(2\pi)^{-s}\Gamma(s+11/2)Z(s)=(2\pi)^{s-1}\Gamma(13/2-s)Z(1-s), \tag{1.7}$$

* Received June 5, 1950; revised January 9, 1952.

[1] The author was a Predoctoral Fellow of the Atomic Energy Commission during the research leading to this paper.

and that $Z(s)$ has infinitely many zeros on the line $\sigma = 1/2$. From (1.6) it follows that $Z(s)$ has no zeros in the half plane $\sigma > 1$; the functional equation, (1.7), then shows that $Z(s)$ has "trivial" zeros at $s = -11/2$, $-13/2, \cdots$ and no other zero with negative real part. Moreover, Rankin [6] has shown that there are no zeros on the lines $\sigma = 0$ and $\sigma = 1$. The "Riemann" hypothesis for $Z(s)$ is that all non-trivial zeros of $Z(s)$ lie on the line $\sigma = 1/2$.

Let G be the (simply connected) region obtained from the half plane $\sigma > 1/2$ by deleting all horizontal segments

$$\sigma + it_0, \qquad 1/2 < \sigma \leq \sigma_0,$$

where $\sigma_0 + it_0$ is a zero, if any exist, of $Z(s)$ in $\sigma > 1/2$. By $\operatorname{Log} Z(s)$ in $\sigma > 1/2$ is meant the analytic continuation into G of $\operatorname{Log} Z(s)$ in $\sigma > 1$, where $\operatorname{Log} Z$ is that branch of the logarithm which is real for real positive Z.

Put

$$Z_n(s) = \prod_{k=1}^{n} (1 - t(p_k) p_k^{-s} + p_k^{-2s})^{-1} \text{ for } \sigma > 1/2.$$

In II it will be shown by a method similar to that used by Bohr in [1], that for every fixed $\sigma > 13/20$, the sequences $\{Z_n(s)\}$ and $\{\operatorname{Log} Z_n(s)\}$ converge in relative measure to $Z(s)$ and $\operatorname{Log} Z(s)$, respectively. In III, the asymptotic distribution functions of Z and of $\operatorname{Log} Z$ will be investigated. (See [5] for definitions of the concepts used in III; much of III is analogous to the corresponding investigation of $\zeta(s)$ in [5].) It will be shown that both asymptotic distribution functions are absolutely continuous with densities of class C^∞, and that for the former, the density vanishes only at $x = 0$ when $13/20 < \sigma < 1$. This is an indication towards the Riemann hypothesis. On the basis of Ramanujan's conjecture (1.2), the same results can be obtained for $\sigma > 1/2$.

II. Convergence in relative measure. In the first part of II, we assume the Ramanujan conjecture. Let $\nu(\sigma) = \text{g. l. b. } \lambda$ for which

$$\int_{-T}^{T} |Z(\sigma + it)|^2 dt = O(T^\lambda), \qquad (T \to \infty).$$

Carlson [3] has shown that $\nu(\sigma)$ is a convex function, and that if

$$q = \text{g. l. b. } \{1 - (1 - v)/\nu(v)\}, \tag{2.1}$$

then

$$\lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} |Z(\sigma + it)|^2 dt$$

exists and is finite for $\sigma > q$. From (1.7) and the asymptotic formula for the Γ-function, it follows that

$$\int_{-T}^{T} |Z(s)|^2 dt = O(T^{2-4\sigma} \int_{-T}^{T} |Z(1-s)|^2 dt) = O(T^{3-4\sigma}),$$

if

$$\int_{-T}^{T} |Z(1-\sigma-it)|^2 dt = O(T). \tag{2.2}$$

When $\sigma < 0$, the Dirichlet series for $Z(1-s)$ is absolutely convergent, so that (2.2) holds. Therefore,

$$\nu(\sigma) \leq 3 - 4\sigma \text{ for } \sigma \leq 0,$$

and in particular $\nu(0) \leq 3$, so that $q \leq 2/3$. Therefore, (2.2) is valid for $\sigma < 1/3$, so that

$$\nu(\sigma) \leq 3 - 4\sigma \text{ for } \sigma \leq 1/3,$$

and $\nu(1/3) \leq 5/3$, which implies that $q \leq 3/5$. On repeating this procedure, one obtains $q \leq (n+1)/(2n+1)$. Letting $n \to \infty$, we have $q \leq 1/2$. But $q < 1/2$ implies, by a theorem of Carlson [2], that

$$\sum_{1}^{\infty} t^2(n) n^{-2\sigma} < \infty, \qquad q < \sigma < 1/2. \tag{2.3}$$

On the other hand,

$$\sum_{1}^{n} t^2(m) \sim an, \qquad a \neq 0$$

(Rankin [7]), so that the series in (2.3) diverges. Hence $q = 1/2$.

If (1.2) is not assumed, then in (2.1), the number 1 must be replaced by 23/20; the corresponding result is that $q \leq 13/20$. Put $\sigma^* = 1/2$ or 13/20 according as (1.2) is or is not true.

By virtue of the definition of $Z_n(s)$,

$$|Z_n(\sigma + it)|^{-1} \leq \prod_{k=1}^{n} (1 + |t(p_k)| p_k^{-\sigma} + p_k^{-2\sigma}) = O(1),$$

as $|t| \to \infty$. Therefore,

$$\int_{-T}^{T} |Z(s)/Z_n(s)|^2 dt = O(T) \text{ as } T \to \infty, \text{ when } \sigma > \sigma^*.$$

From the inequality $(a-1)^2 \leq 2a^2 + 2$, it follows that

$$\int_{-T}^{T} |Z(s)/Z_n(s) - 1|^2 dt = O(T), \text{ when } \sigma > \sigma^*;$$

so that by the theorem of Carlson in [3] referred to above,

$$\lim_{T\to\infty}\int_{-T}^{T}|\,Z(s)/Z_n(s)-1\,|^2dt=\sum_{m=1}^{\infty}c^2{}_n(m)m^{-2\sigma}\quad\text{for }\sigma>\sigma^*,$$

where $\sum_{m=1}^{\infty}c_n(m)m^{-s}=Z(s)/Z_n(s)-1$ for $\sigma>1$, say. But for $\sigma>1$,

$$Z(s)/Z_n(s)=\prod_{k=n+1}^{\infty}(1-t(p_k)p_k^{-s}+p_k^{-2s})^{-1},$$

so that

$$c_n(1)=c_n(2)=\cdots=c_n(n)=0,$$

while $|\,c_n(m)\,|\leq|\,t(m)\,|$ for all m. Thus, for any fixed $\sigma>\sigma^*$,

$$\lim_{T\to\infty}\tfrac{1}{2T}\int_{-T}^{T}|\,Z(s)/Z_n(s)-1\,|^2dt\leq\sum_{n+1}^{\infty}t^2(m)m^{-2\sigma}\to 0\text{ as }n\to\infty.$$

Therefore, if meas $_T\{|\,Z(s)/Z_n(s)-1\,|>\epsilon\}$ denotes the Lebesgue measure of the set of t- vaules for which $-T\leq t\leq T$ and $Z(s)/Z_n(s)-1\,|>\epsilon$, then

$$\lim_{n\to\infty}\limsup_{T\to\infty}\text{ meas }_T\{|Z(s)/Z_n(s)-1\,|>\epsilon\}/2T=0.$$

Since $|\,\text{Log}\,Z-\text{Log}\,Z_n\,|=|\,\text{Log}\,[1+(Z/Z_n-1)]\,|<|\,Z/Z_n-1\,|$, when $|\,Z/Z_n-1\,|$ is sufficiently small, it follows that $\{\text{Log}\,Z_n(s)\}$ converges in relative measure to $\text{Log}\,Z(s)$ for every fixed $\sigma>\sigma^*$; furthermore, since $Z_n(s)=O(1)$ as $|\,t\,|\to\infty$, it follows as before, that

$$\lim_{T\to\infty}\tfrac{1}{2T}\int_{-T}^{T}|\,Z(s)-Z_n(s)\,|^2dt\leq\sum_{n+1}^{\infty}4t^2(m)m^{-2\sigma},\qquad\sigma>\sigma^*;$$

so that for every fixed $\sigma>\sigma^*$, $\{Z_n(s)\}$ converges in relative measure to $Z(s)$.

III. The distribution functions. We shall prove

THEOREM 3.1. *Put*

$$F_k(z)=-\text{Log}(1-t(p_k)z+z^2);$$

for fixed σ, let S_k denote the curve

$$x_k(\theta)=\xi_k(\theta)+i\eta_k(\theta)=F_k(p_k^{-\sigma}e^{2\pi i\theta}),$$

and let ϕ_k be the distribution function determined by this parametric representation of S_k; that is, of E is a Borel set in the x-plane, then $\phi_k(E)$ is the θ measure of ES_k. Let $\Lambda(y,\phi_k)$ denote the Fourier transform of ϕ_k:

$$\Lambda(y,\phi_k)=\int_0^1\exp(ig_{\tau,k}(\theta)\,|\,y\,|)\,d\theta,$$

where $y = (|y| \cos 2\pi\tau, |y| \sin 2\pi\tau)$ *and* $g_{\tau,k} = \xi_k(\theta) \cos 2\pi\tau + \eta_k(\theta) \sin 2\pi\tau$. *Then, for* $\sigma > \sigma^*$,

$$\Lambda(y, \phi_k) = O(|y|^{-\frac{1}{2}}) \text{ as } |y| \to \infty,$$

for infinitely many k.

Proof. For convenience we drop the subscript k. We apply Theorem 12 of [5]. It is clear that ξ and η possess continuous second derivatives; we must show that S is convex and that $g_\tau''(\theta)$ has, for every fixed τ, exactly two zeros θ on S.

An application of the Study Rundungsschranke ([8], p. 109) gives the result that a sufficient condition that S be convex is that

$$13p^{-\sigma} \leq |t(p)| \leq p^\sigma/2 \text{ and } p \text{ is sufficiently large.} \tag{3.1}$$

In fact, the Study theorem states that S is convex if $F'(0) \neq 0$ and

$$\Re[1 + zF''/F'] > 0 \text{ for } |z| = p^{-\sigma}.$$

Now $F'(0) = t(p)$, while

$$\begin{aligned} |(t(p) - 2z)(1 - t(p)z + z^2)|^2 \Re[1 + zF''/F'] = \\ (t^2(p) + 3t^2(p)p^{-4\sigma} + 8p^{-2\sigma} - 8p^{-4\sigma}) \\ - (t^3(p) + 6t(p) + t^3(p)p^{-2\sigma} + 8t(p)p^{-2\sigma} + 2t(p)p^{-4\sigma})p^{-\sigma} \cos 2\pi\theta \\ + (8t^2(p) + 16p^{-2\sigma})p^{-2\sigma} \cos^2 2\pi\theta - 8t(p)p^{-3\sigma} \cos^3 2\pi\theta, \end{aligned}$$

where $z = p^{-\sigma}e^{2\pi i\theta}$. Clearly it is sufficient to consider $t(p) > 0$ and $\cos 2\pi\theta > 0$. The last two terms of the preceding expression can be written in the form

$$8p^{-2\sigma}(t^2(p) + 2p^{-2\sigma} - t(p)p^{-\sigma} \cos 2\pi\theta) \cos^2 2\pi\theta.$$

This is non-negative, since $2|t(p)|p^{-\sigma} \leq t^2(p) + p^{-2\sigma}$. The last two members, $8p^{-2\sigma} - 8p^{-4\sigma}$, of the term independent of θ dominate the last two members of the coefficient of $\cos 2\pi\theta$,

$$8p^{-2\sigma} - 8p^{-4\sigma} \geq 8t(p)p^{-3\sigma} + 2t(p)p^{-5\sigma} \text{ if } p \geq 3,$$

by virtue of the right hand inequality of (3.1). Similarly, the first member, $t^2(p)$, of the term independent of θ dominates the first three members of the coefficient of $\cos 2\pi\theta$, for

$$t^2(p) - (t^3(p)p^{-\sigma} + 6t(p)p^{-\sigma} + t^3(p)p^{-3\sigma}) \geq t^2(p)/2 - 25t(p)p^{-\sigma}/4 \geq 0$$

follows from (3.1). Thus (3.1) is sufficient to assure the convexity of S; by a more tedious computation it may be shown that if $0 < |t(p)| \leq p^\sigma$,

and if $2p^{-\sigma} \leq |t(p)| \leq 4p^{-\sigma}$ *does not hold*, then S is convex. The right hand inequality of (3.1) is superfluous in view of (1.1) and $\sigma > \sigma^* \geq 1/2$; also (3.1) may be replaced by the condition

$$|t(p)| \geq p^{-1/2} \text{ and } p \text{ sufficiently large.} \tag{3.2}$$

That some such condition is necessary is shown by the fact that if $t(p) = 3p^{-\sigma}$ and $z = p^{-\sigma}$, then $\Re[1 + zF''/F'] < 0$.

Infinitely many primes satisfy (3.2), for if not, the series (1.5) would converge absolutely for $\sigma > 1/2$. It remains to show that for those primes satisfying (3.2), $g_\tau''(\theta)$ has just two zeros. Put $\nu = \nu_k = -1$ if $t(p) < -2$, and $\nu = 1$ otherwise. If $\gamma = \gamma_k$ is the real or purely imaginary angle satisfying

$$2\cos 2\pi\gamma = \nu t(p) \text{ and } 0 \leq \gamma < 1/2 \text{ or } i\gamma < 0,$$

then

$$F(z) = 2\sum_1^\infty \nu^n z^n \cos 2\pi n\gamma/n, \qquad |z| \leq p^{-1/2}. \tag{3.3}$$

Furthermore, it is easy to see that $F(p^{-\sigma}e^{2\pi i\theta}) = -\operatorname{Log}(1-\zeta_1)(1-\zeta_2)$, where $\zeta_1 = \nu p^{-\sigma}e^{2\pi i(\theta-\gamma)}$, $\zeta_2 = \nu p^{-\sigma}e^{2\pi i(\theta-\gamma)}$; so that

$$g_\tau''(\theta) = -4\pi^2\Re\{[\zeta_1(1-\zeta_1)^{-2} + \zeta_2(1-\zeta_2)^{-2}]e^{-2\pi i\tau}.$$

If $p \to \infty$ through the primes satisfying (3.2), then

$$\zeta_1,\ \zeta_2,\ (\zeta_2 e^{2\pi i\gamma} + \zeta_1 e^{-2\pi i\gamma})/\cos 2\pi\gamma,\ \ (\zeta_2^2 e^{2\pi i\gamma} - \zeta_1^2 e^{-2\pi i\gamma})/\cos 2\pi\gamma$$

tend to 0 uniformly for all θ, by virtue of (1.1) and $\sigma > \sigma^*$. Therefore,

$$g_\tau''(\theta) = -8\pi^2\nu p^{-\sigma}\cos 2\pi\gamma \cos 2\pi(\theta-\tau) + o(1)$$

uniformly for all θ and τ, as $p \to \infty$ through those primes satisfying (3.2). Hence, the zeros of $[g_\tau''(\theta)/\nu p^{-\sigma}\cos 2\pi\gamma]$ cluster only at $\theta = \tau \pm 1/4$, the zeros of $\cos 2\pi(\theta-\tau)$. (It is understood that p is restricted to those primes satisfying (3.2).) In the same fashion it may be shown that the zeros of g''' are near $\theta = \tau,\ \tau + 1/2$. Since there is a zero of g''' between every two zeros of g'', it follows that if p is a sufficiently large prime satisfying (3.2), there is at most one zero of g'' near each of $\theta = \tau \pm 1/4$; inasmuch as a continuous derivative of a continuous periodic function has at least two zeros, it is clear that for all primes satisfying (3.2), $g_\tau''(\theta)$ has for every fixed τ precisely two zeros θ on S, which completes the proof of Theorem 3.1.

Next we show that for any fixed $\sigma > \sigma^*$, the infinite convolution $\phi_1 * \phi_2 \cdots$ is absolutely convergent. Now $x_k(\theta) \to 0$ uniformly for all θ, as $k \to \infty$; hence, S_k, the spectrum of ϕ_k, is contained in a fixed circle. If $c(\phi_k)$ and

$M(\phi_k)$ are the first and second moments of ϕ_k, respectively, then $c(\phi_k) = 0$, so that the convergence of

$$\sum_1^\infty M(\phi_k) = 4 \sum_1^\infty \sum_1^\infty p_k^{-2m\sigma} \cos^2 2\pi m\gamma_k/m^2$$

is necessary and sufficient for the convergence and absolute convergence of the infinite convolution $\phi = \phi_1 * \phi_2 * \cdots$, by Theorems 5 and 6 of [5]. The convergence of the double series is obvious if Ramanujan's hypothesis is true. Otherwise, $\cos^2 2\pi m\gamma_k = O(p_k^{3m/5})$ as $m \to \infty$, uniformly in k. Hence

$$\sum_{k=1}^\infty \sum_{m=2}^\infty p_k^{-2m\sigma} \cos^2 2\pi m\gamma_k/m^2 = O(1) \sum_{k=1}^\infty \sum_{m=2}^\infty p_k^{m(3/5-2\sigma)}/m^2,$$

the last series being convergent for $\sigma > 11/20$; also

$$\sum_1^\infty p_k^{-2\sigma} \cos^2 2\pi\gamma_k = O(1) \sum t^2(p) p^{-2\sigma},$$

which converges for $\sigma > 13/20$. Thus, the infinite convolution is absolutely convergent for $\sigma > \sigma^*$.

Since $|\Lambda(y, \phi_k)| \leq 1$ for all k, and $\Lambda(y, \phi_k) = O(|y|^{-\frac{1}{2}})$ for infinitely many k, we have $\Lambda(y, \phi) = \Lambda(y, \phi_1)\Lambda(y, \phi_2) \cdots = O(|y|^{-m})$ as $|y| \to \infty$, for every fixed m. Therefore (cf. [5], § 3), ϕ is absolutely continuous, and its density $D(x)$ has continuous partial derivatives of all orders.

There is an essential difference in the spectrum of ϕ if $\sigma > 1$ and if $1 > \sigma > \sigma^*$. In the former case, the spectrum is a bounded set. This follows from Theorem 3 of [5]. On the other hand, it will now be shown that if $1 > \sigma > \sigma^*$, the spectrum of ϕ is the entire plane. First $\sum |t(p)| p^{-\sigma} = \infty$, for if this were false it would follow from the product representation (1. 6) of $Z(s)$ that $\sum |t(n)| n^{-\sigma} < \infty$, which is false. The following calculation of the spectrum for $1 > \sigma > \sigma^*$ is modelled on the proof of the second part of Theorem 14 of [5]. Let $1 > \sigma > \sigma^*$ be fixed. If C is a sufficiently large constant, then

$$|F_k(z) - t(p_k)z| \leq C \max(4, t^2(p_k))|z|^2 \text{ for } |z| = p_k^{-\sigma}.$$

This follows from the representation (3. 3) of $F_k(z)$, since

$$|\cos 2\pi m\gamma| \leq |e^{-2\pi i m\gamma}|$$

implies that

$$|F_k(z) - t(p_k)z| \leq 2|e^{-2\pi i\gamma z}|^2/(1 - |e^{-2\pi i\gamma z}|).$$

Clearly, $\sum \max(4, t^2(p_k)) p_k^{-2\sigma} < \infty$, since both $\sum p_k^{-2\sigma}$ and $\sum t^2(p_k) p_k^{-2\sigma}$ converge. Let $\epsilon > 0$ be arbitrary, and fix m so large that

$$C \sum_{m+1}^{\infty} \max(4, t^2(p_k)) p_k^{-2\sigma} < \epsilon.$$

Let x_0 be any point of the x-plane, and let x_1 be an arbitrary point of the form

$$F_1(p_1^{-\sigma} e^{2\pi i \theta_1}) + \cdots + F_m(p_m^{-\sigma} e^{2\pi i \theta_m}).$$

Then, if $n > m$ is large enough, $\sum_{m+1}^{n} | t(p_k) | p_k^{-\sigma} > | x_0 - x_1 |$, and none of the numbers $| t(p_k) | p_k^{-\sigma}$, $k = m+1, m+2, \cdots, n$ is greater than the sum of the others. Therefore, we may choose $\theta_{m+1}, \theta_{m+2}, \cdots, \theta_n$ so that $\sum_{m+1}^{n} t(p_k) p_k^{-\sigma} e^{2\pi i \theta_k} = x_0 - x_1$, which implies that $| \sum_{1}^{n} F_k(p_k^{-\sigma} e^{2\pi i \theta_k}) - x_0 | < \epsilon$, so that the spectrum of ϕ is the entire plane.

It is true, moverover, that for any σ satisfying $1 > \sigma > \sigma^*$, the density of ϕ is positive for all x. For a proof of this, cf. [5], p. 60.

Remark. The same will be true for $\sigma = 1$ if $\sum | t(n) | / n = \infty$. If the latter series converges, the spectrum of ϕ for $\sigma = 1$ is a bounded set. We may sum up the preceding remarks in the following theorem.

THEOREM 3.2. *For any fixed $\sigma > \sigma^*$, the infinite convolution $\phi_1 * \phi_2 * \cdots$ is absolutely convergent, $\phi = \phi_\sigma = \phi_1 * \phi_2 * \cdots$ is absolutely continuous, and its density $D_\sigma(x)$ possesses continuous partial derivatives of all orders. If $1 > \sigma > \sigma^*$, then $D_\sigma(x) > 0$ for all x. The spectrum is symmetric with respect to the real axis; the spectrum of ϕ_σ is a bounded domain if $\sigma > 1$, and is the whole x-plane if $1 > \sigma > \sigma^*$. For any fixed $\lambda > 0$, $D_\sigma(x) = O(e^{-\lambda |x|^2})$ as $| x | \to \infty$; this appraisal is valid for each partial derivative of $D_\sigma(x)$.*

The syymetry of the spectrum is obvious; a proof of the last assertion may be obtained by modifying the proof of Theorem 16 of [5] in the same way that the proof of part of Theorem 14 was modified to verify that for $1 > \sigma > \sigma^*$, the spectrum of ϕ is the whole plane.

Still following [5] we define $\tilde{\phi}_\sigma$ by $\tilde{\phi}_\sigma(E) = \phi_\sigma(\log E)$; the following may readily be obtained:

THEOREM 3.3. *The distribution $\tilde{\phi}_\sigma$ is symmetric with respect to the real axis; if $1 > \sigma > \sigma^*$, its spectrum is the entire plane; if $\sigma > 1$, the spectrum is a bounded set not containing the origin. This distribution is absolutely continuous with a density $\tilde{D}_\sigma(x)$ which possesses continuous partial derivatives of all orders. For any fixed $\lambda > 0$, $\tilde{D}_\sigma(x) = O(e^{-\lambda \log^2 |x|})$ as $| x | \to \infty$ or $| x | \to 0$; these appraisals are valid also for each partial derivative of $\tilde{D}_\sigma(x)$. If $1 > \sigma > \sigma^*$, then $\tilde{D}_\sigma(x) = 0$ if and only if $x = 0$.*

Finally, we have the analogues of Theorems 30 and 31 of [5]:

THEOREM 3.4. *The functions* $\operatorname{Log} Z(\sigma+it)$ *and* $Z(\sigma+it)$ *possess asymptotic distribution functions in the unrestricted case if* $\sigma>1$, *and in the restricted case if* $1\geq\sigma>\sigma^*$. *These are, respectively, the distribution functions* ϕ_σ *and* $\bar{\phi}_\sigma$ *discussed above. The spectrum of* ϕ_σ *is the closure of the range of* $\operatorname{Log} Z(\sigma+it)$, *and the spectrum of* ϕ_σ *is the closure of the range of* $Z(\sigma+it)$ *for* $-\infty<t<\infty$.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] H. Bohr, "Zur Theorie der Riemann'schen Zetafunktion im kritischen Streifen," *Acta Mathematica*, vol. 40 (1916), pp. 67-100.

[2] F. Carlson, "Contributions à la théorie des séries de Dirichlet I," *Arkiv för Matematik Astronomi och Fysik*, vol. 16 (1922), No. 18.

[3] ———, "Contributions à la théorie des séries de Dirichlet II," *Arkiv för Matematik Astronomi och Fysik*, vol. 19 (1926), No. 25.

[4] G. Hardy, "*Ramanujan*," Cambridge University Press, 1940.

[5] B. Jessen and A. Wintner, "Distribution functions and the Riemann zeta function," *Transactions of the American Mathematical Society*, vol. 38 (1935), pp. 48-88.

[6] R. Rankin, "Contributions to the theory of Ramanujan's function $\tau(n)$ and similar arithmetical functions I," *Proceedings of the Cambridge Philosophical Society*, vol. 35 (1939), pp. 351-356.

[7] ———, "Contributions to the theory of Ramanujan's function $\tau(n)$ and similar arithmetical functions II," *Proceedings of the Cambridge Philosophical Society*, vol. 35 (1939), pp. 357-372.

[8] E. Study, "*Vorlesungen über ausgewählte Gegenstände der Geometrie*," B. G. Teubner, 1913.

[9] J. Wilton, "A note on Ramanujan's arithmetical function $\tau(n)$," *Proceedings of the Cambridge Philosophical Society*, vol. 25 (1929), pp. 121-129.

ANALYTIC FUNCTIONS OF BOUNDED TYPE.*

By A. FRANCK.

1. Introduction. An analytic function $f(z)$ is called *of bounded type* in a region if it can be represented as a quotient of two bounded analytic functions in that region, i. e. $f(z) = h_1(z)/h_2(z)$ where $|h_i(z)| < 1$, $(i = 1, 2)$. The following theorem is fundamental for the characterization of the class of functions of bounded type which are analytic in the unit circle.

OSTROWSKI-NEVANLINNA THEOREM.[1] (i) *A necessary and sufficient condition that a function $F(w)$, regular in the circle $|w| < 1$, be* B. T. *in $|w| < 1$ is that*

$$\int_0^{2\pi} \log^+ |F(re^{i\theta})|\, d\theta \leqq M, \qquad 0 \leqq r < 1 \tag{1.1}$$

where M is a constant.[2] (ii) *If $F(w)$ is* B. T. *in $|w| < 1$, then $F(w)$ may be represented in the form $F(w) = B(w)G(w)$ where $B(w)$ is the Blaschke product formed with the zeros $w_1, w_2, \cdots$ of $F(w)$ in $|w| < 1$, i. e.* $B(w) = \prod_{k=1}^{\infty} \frac{w - w_k}{\bar{w}_k w - 1} \frac{\bar{w}_k}{|w_k|}$, *the product being uniformly and absolutely convergent in every circle $|w| \leqq \rho < 1$; $G(w)$ is a non-vanishing function* B. T. *in the unit circle and $\log |G(w)|$ is representable by a Poisson-Stieltjes integral in $|w| < 1$*

$$\log |G(w)| = \tfrac{1}{2\pi} \int_0^{2\pi} \frac{1 - \rho^2}{1 + \rho^2 - 2\rho \cos(\theta - \alpha)}\, d\phi(\theta), \tag{1.2}$$

where $w = \rho e^{i\alpha}$ and $\phi(\theta)$ is a function of bounded variation in $0 \leqq \theta \leqq 2\pi$. (iii) *If the inequality* (1.1) *is satisfied for $F(w)$, then it is also satisfied for $G(w)$ [with the same constant M].* (iv) *If $F(w)$ is* B. T. *in $|w| < 1$, it*

---

* Received March 8, 1950; revised May 4, 1951; presented to the American Mathematical Society, November 26, 1949. This paper contains a part of the results of the author's doctoral dissertation [1]. The author wishes to express his indebtedness to Professor S. E. Warschawski, under whose direction this work was done.

[1] Hereafter this theorem will be referred to as the O. N. Theorem. Furthermore a function of bounded type will be denoted by B. T.

[2] $\log^+ a = \log a$ for $a \geqq 1$ and $\log^+ a = 0$ for $a < 1$.

possesses boundary values for almost all θ, $(w = e^{i\theta})$ *on the unit circle for non-tangential approach.*

These results are due to A. Ostrowski [6]; condition (1. 1) and the result on the existence of the boundary values were also given in a joint paper by F. and R. Nevanlinna [4], p. 26.

The concept of a function of B. T. can be applied to a function which is meromorphic (rather than just regular) in a region. In a subsequent paper [5], R. Nevanlinna extended part (i) of the above theorem not only to functions meromorphic in the unit circle, but also to functions meromorphic in an arbitrary region. In the present paper we shall restrict ourselves to analytic functions.

The criteria obtained by R. Nevanlinna were based on the use of the Green's function for the regions considered. It would seem desirable to obtain criteria which are not based on a knowledge of the Green's function, which in general is difficult to obtain, but rather are expressed in terms of the geometry of the region of analyticity of the functions considered.

The purpose of the present paper is to establish such criteria for a fairly general class of regions. Applications of our methods permit us to derive several results which hold for special regions (such as the half-plane). In particular our criteria contain some of the results of A. Wishard [11] for the case of analytic functions.[3]

I. General criteria.

2. Extensions of the O. N. Theorem for the Half-Plane. The following lemma is an analogue of the O. N. Theorem stated for the half-plane.

LEMMA 1. *Suppose that* $f(z)$ *is analytic in* $\mathrm{Re}(z) > 0$, $z = x + iy$. *Let* C_n *be a sequence of circles, radii* ρ_n, *which exhaust*[4] $x > 0$, $(n = 1, 2, \cdots)$. *Suppose that for all* n

---

[3] Other classes of functions which are regular in a half-plane were treated by W. Kryloff [3]. He studies the classes of functions for which $(z = x + iy)$ (a) $\int_{-\infty}^{\infty} |f(z)|^p dy \leqq M^p < \infty$ and (b) $\int_{-\infty}^{\infty} \log^+ |f(z)|\, dy \leqq M < \infty$ and obtains "parametric representations" (in terms of certain Stieltjes integrals) for these functions.

[4] *A sequence of closed Jordan* (i. e. simple) *curves* C_n which are contained in a region R is said to exhaust the region R, if for any subregion R' whose closure R' is contained in R, there exists an integer N such that R' will be contained in the interior of C_n for all $n > N$. A similar definition is to be applied to the concept that *a sequence of regions exhaust a given region.*

$$\int_{C_n} \frac{\log^+ |f(z)|}{|1+z|^2} \rho_n \, d\theta \leqq M.^5 \tag{2.1}$$

Then $f(z)$ *is* B. T. *in* $x > 0$; $f(z) = b(z)g(z)$ *where* (i) $b(z)$ *is the Blaschke product formed with the zeros* $z_1, z_2, \cdots$ *of* $f(z)$ *in* $x > 0$; (ii) $g(z) \neq 0$; (iii) *for* $x > 0$ *there exists a constant* λ *such that*

$$\log |g(z)| = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x \, d\mu(t)}{x^2 + (y-t)^2} + \lambda x \tag{2.2}$$

where $\mu(t)$ *is a function of bounded variation in any finite interval and the integral converges absolutely*; (iv) *there exists a constant* N *such that for any circle* C *of radius* ρ *contained in* $x > 0$

$$\int_C \frac{\log^+ |g(z)|}{|1+z|^2} \rho \, d\theta \leqq N. \tag{2.3}$$

Conversely if $f(z)$ *is* B. T. *in* $x > 0$, *then it satisfies* (2.3) *with* f *replacing* g.

Proof. Our proof makes use of a theorem of R. Gabriel [2] for functions $U(w)$ non-negative and subharmonic in the unit circle. Let C_1 and C_2 *be circles contained in* $|w| < 1$ *such that* C_2 *lies in the interior* of C_1. *Then*

$$\int_{C_2} U(w) \, |dw| \leqq 2 \int_{C_1} U(w) \, |dw|.$$

We map $x > 0$ onto $|w| < 1$ by means of the transformation

$$w = \frac{z-1}{z+1} \tag{2.4}$$

carrying the circles C_n onto a sequence of circles Γ_n which exhaust $|w| < 1$. Points on C_n are given by $z = z_n + \rho_n e^{i\theta}$ where z_n is the center of C_n. Employing the inverse of (2.4) and $|dw| = 2\rho_n d\theta / |1+z|^2$, we find that

$$\int_{C_n} \frac{\log^+ |f(z)|}{|1+z|^2} \rho_n \, d\theta = \tfrac{1}{2} \int_{\Gamma_n} \log^+ |F(w)| \, |dw| \tag{2.5}$$

where $F(w) = f[(1+w)/(1-w)]$. Since the left side of (2.5) $\leqq M$. we have

$$\int_{\Gamma_n} \log^+ |F(w)| \, |dw| \leqq 2M. \tag{2.6}$$

Consider the circle $|w| = r_0 < 1$. Keeping r_0 fixed, choose n so large that Γ_n will contain this circle in its interior. Applying Gabriel's Theorem

[5] As a matter of convention the letters $M, M_1, M_2, \ldots, N, N_1, N_2, \ldots$ will always designate constants.

to $\log^+ |F(w)|$, we find from (2.6) that $\int_0^{2\pi} \log^+ |F(r_0 e^{i\phi})|\, d\phi \leqq 4M/r_0$. Since $\int_0^{2\pi} \log^+ |F(re^{i\phi})|\, d\phi$ is a monotonic increasing function of r, we have for any circle $|w| = r < r_0$: $\int_0^{2\pi} \log^+ |F(re^{i\phi})|\, d\phi \leqq 4M/r_0$. Keeping r fixed, let $r_0 \to 1$. Thus $\int_0^{2\pi} \log^+ |F(re^{i\phi})|\, d\phi \leqq 4M$ for all r, $0 \leqq r < 1$. Hence by the O. N. Theorem $F(w)$ is B. T. in $|w| < 1$; $F(w) = B(w)G(w)$ where $G(w) \neq 0$. Returning to $x > 0$, we find $f(z) = b(z)g(z)$ where $g(z) = G[(z-1)/(z+1)] \neq 0$. By means of the mapping (2.4) and (1.2) of the O. N. Theorem, the representation for $\log |g(z)|$ is established.

To establish (2.3) we map $x > 0$ onto $|w| < 1$ carrying a circle C in $x > 0$ into a circle Γ in $|w| < 1$. Let r_0 be the maximal distance of Γ from $w = 0$. Construct any circle $|w| = r > r_0$, $0 \leqq r < 1$. Applying Gabriel's Theorem to $\log^+ |G(w)|$ we have

$$\int_\Gamma \log^+ |G(w)|\, |dw| \leqq 2 \int_0^{2\pi} \log^+ |G(re^{i\phi})|\, r\, d\phi. \tag{2.7}$$

From the O. N. Theorem the integral on the right side of (2.7) is bounded for all r, $0 \leqq r < 1$. Transforming the left side of (2.7) by (2.4), we establish (2.3).

To prove the converse merely repeat for $f(z) = f[(1+w)/(1-w)] = F(w)$ the argument which we applied to $g(z) = G(w)$ in the preceding paragraph.

The following theorem is readily established from Lemma 1.

THEOREM I. *Let $f(z)$ be analytic in $x > 0$. Let C_n be a sequence of circles, radii R_n, which exhaust $x > 0$. Suppose that for all n*

$$\iint_{C_n} \frac{\log^+ |f(z)|}{|1+z|^2}\, dx\, dy \leqq MR_n. \tag{2.8}$$

Then $f(z)$ is B. T. *in $x > 0$; $f(z) = b(z)g(z)$ and for any circle C in $x > 0$*

$$\iint_C \frac{\log^+ |g(z)|}{|1+z|^2}\, dx\, dy \leqq NR, \qquad (R = \text{radius of } C). \tag{2.9}$$

Conversely, if $f(z)$ is B. T. *in $x > 0$. then $f(z)$ satisfies (2.9) with f replacing g.*

Proof. Consider the annular rings A_{R_n}: $(R_n/2 \leqq |z - z_n| \leqq R_n)$ where z_n is the center of C_n. Employing polar coordinates about $z = z_n$, we find from (2.8) that

$$\int_{R_n/2}^{R_n} \left[\int_0^{2\pi} \frac{\log^+ |f(z)|}{|1+z|^2} d\theta \right] r\, dr \leqq MR_n. \tag{2.10}$$

Applying the mean value theorem to the outer integral of (2.10), we have

$$\iint_{A_{R_n}} \frac{\log^+ |f(z)|}{|1+z|^2} d\theta\, r\, dr = \frac{3R^2_n}{8} \int_0^{2\pi} \frac{\log^+ |f(\zeta)|}{|1+\zeta|^2} d\theta \tag{2.11}$$

where $\zeta = z_n + \rho_n e^{i\theta}$ and $R_n/2 < \rho_n < R_n$.

Since $\rho_n < R_n$, we have from (2.10) and (2.11)

$$\int_0^{2\pi} \frac{\log^+ |f(\zeta)|}{|1+\zeta|^2} d\theta \leqq \frac{8M}{3\rho_n}.$$

Thus the conditions of Lemma 1 are satisfied for the circles $|z - z_n| = \rho_n$ and hence $f(z)$ is B. T. in $x > 0$. From Lemma 1 it follows that for *any* circle C in $x > 0$, we have $\int_C \frac{\log^+ |g(z)|}{|1+z|^2} \rho\, d\theta \leqq N$. From this, inequality (2.9) follows. The converse also follows from the converse of Lemma 1.

3. An area theorem for arbitrary regions. In the present section we shall extend our results to obtain conditions under which an analytic function will be B. T. in a region [6] of a more general nature than a circle or a half-plane. The following theorem which is an extension of Theorem I gives a sufficient condition for an analytic function to be B. T. in a region R.

THEOREM II. (a) *Let $f(z)$ be analytic in a bounded region R. Let R_n be a sequence of regions which exhaust R. Let $w = \phi(z)$ map R conformally onto $|w| < 1$ so that a point $z = a$ on the boundary of R correspond to $w = 1$, suppose that* (b) $|\phi'(z)| \leqq m$ *in R where m is a constant and* (c) $\left|\frac{1-w}{a-z}\right| \geqq N > 0$ *in R, where $w = \phi(z)$.*[7] (d) Let $d_n = \sup d_n(z)$ *for all z on the boundary of R_n where $d_n(z)$ is the minimal distance of the point z on the boundary of R_n to the boundary of R.*

---

[6] By region we mean a simply-connected region. This usage will be applied throughout the remainder of the paper.

[7] Hypotheses (b) and (c) of this theorem will, for example, be satisfied if R is bounded by a closed Jordan curve C, with the following properties: (i) one can roll a circle on the exterior of C (i. e. to each point ζ of C, one can associate a circle passing through ζ such that its interior lies in the exterior of C, each circle having the same radius); (ii) at the point $z = a$, the curve possesses finite curvature (by this we mean that C possesses a tangent at every point of a neighborhood of $z = a$ and that the curvature exists and is finite at $z = a$).

Conditions (i) implies (b), for example, see [9]; condition (ii) implies (c), for example, see [10], p. 433 or [8], Theorems 18 and 19.

Then $f(z)$ is B. T. *in R if*

$$\iint_{R_n} \frac{\log^+ | f(z)|}{| a - z |^2}\, dx\, dy \leqq \frac{M}{d_n} \tag{3.1}$$

Proof. Since $w = \phi(z)$ maps R onto $| w | < 1$, it also maps the regions R_n onto regions W_n contained in $| w | < 1$ which exhaust the unit circle. If n is taken sufficiently large, the regions W_n will contain the point $w = 0$. Consider such regions which contain $w = 0$ and let r_n be the minimal distance of the boundary of W_n from $w = 0$. Let us consider a point on the boundary of W_n for which the minimal distance is assumed. Designate it as $w = r_n e^{i a_n}$. Let z_n be the image point in the z-plane as defined by $r_n e^{i a_n} = \phi(z_n)$. Consider the integral $\int_{z_n}^{z_0} | \phi'(z)| \, | dz |$ along the straight line segment joining z_n to the nearest point z_0 on the boundary of R. Transforming this integral to the w-plane by means of the mapping function, we have

$$\int_{z_n}^{z_0} | \phi'(z)| \, | dz | = \int_{r_n e^{i a_n}}^{e^{i\beta}} | dw |, \qquad [e^{i\beta} = \phi(z_0)]; \tag{3.2}$$

the last integral being taken over the image curve of the line segment from z_n to z_0. From hypothesis (b) and the fact that the length of any curve from any point w in $| w | < 1$ to the boundary is greater than or equal to $1 - | w |$, we find from (3.2) that $m d_n(z_n) \geqq 1 - r_n$ where $d_n(z_n)$ is the minimal distance from the point z_n to the boundary of R. Thus, we finally have that

$$d_n \geqq (1 - r_n)/m. \tag{3.3}$$

Now let $F(w) = f[\psi(w)]$ where $z = \psi(w)$ is the inverse of $w = \phi(z)$. On transforming the integral in (3.1) to the w-plane. we find from (b) and (c) that

$$\iint_{R_n} \frac{\log^+ | f(z)|}{| a - z |^2}\, dx\, dy = \iint_{W_n} \frac{\log^+ | F(w)|}{| 1 - w |^2} \, \frac{| 1 - w |^2}{| a - z |^2} \, | \psi'(w) |^2 \, du\, dv$$

$$\geqq \frac{N^2}{m^2} \iint_{W_n} \frac{\log^+ | F(w)|}{| 1 - w |^2}\, du\, dv$$

as $| \psi'(w) | = 1/| \phi'(z) | \geqq 1/m$. Taking note of (3.1) and (3.3) we find that

$$\iint_{W_n} \log^+ | F(w)|/| 1 - w |^2 \, du\, dv \leqq (m^3 M/N^2)/(1 - r_n).$$

By means of the mapping $w = (\zeta - 1)/(\zeta + 1)$ to $\mathrm{Re}(\zeta) \geqq 0$, one can show that condition (2.8) is satisfied. Hence it follows by Theorem I that $f(z) = F[\phi(z)]$ is B. T. in R.

As a converse to this theorem we establish the following corollary.

COROLLARY. *Let $f(z)$ be* B. T. *in a bounded region R. Let $w=\phi(z)$ map the region R onto the unit circle $|w|<1$ in such a manner that a point $z=a$ on the boundary of R corresponds to $w=1$. Suppose that $|\phi'(z)| \geqq m_1 > 0$ in R where m_1 is a constant and that $\left|\frac{1-w}{a-z}\right| \leqq N_1$ in R where $w=\phi(z)$.*

Then there exists a sequence of regions R_n which exhaust R such that

$$\iint_{R_n} [\log^+ |f(z)|/|a-z|^2]\,dx\,dy \leqq N/d_n \tag{3.4}$$

where d_n has the same meaning as in Theorem II.

Proof. Consider a sequence of circles Γ_n in $|w|<1$ which exhaust the unit circle. By employing an argument similar to that of the proof of Theorem II, one can show that the images, R_n, of the interior of the Γ_n form a sequence for which (3.4) is satisfied.

REMARK. *If the region R is such that there exists two positive constants m_1 and m_2 for which $m_1 \leqq |\phi'(z)| \leqq m_2$,[8] then condition* (3.1) *is necessary and sufficient for $f(z)$ to be* B. T. *in R.*

The following theorem can also be deduced from the O. N. Theorem. It is merely stated here. For a proof of this results, see [1].

THEOREM III. *Let $f(z)$ be analytic in the interior R of a closed Jordan curve. Suppose that $w=\phi(z)$ maps R conformally onto $|w|<1$ so that $0<m_1 \leqq |\phi'(z)| \leqq m_2$ [8] for all z in R where m_1 and m_2 are constants. Then a necessary and sufficient condition for $f(z)$ to be* B. T. *in R is that there exist some sequences of regions R_n which exhaust R, for which*

$$\iint_{R_n} [\log^+ |f(z)|/d(z)]\,dx\,dy \leqq M \log 1/d_n$$

where $d(z)$ is the distance of the point z to the boundary of R and d_n has the same meaning as in Theorem II.

II. Applications.

4. A criterion for a circular sector. As an application of the general area Theorem II, we establish

THEOREM IV. *Let $f(z)$ be analytic in the sector*

$$S: (0<r<1, -\alpha<\theta<\alpha).$$

[8] This condition can be characterized in terms of the geometry of the boundary of R (see [8], Theorems 18 and 19 and [10], p. 433).

Suppose that $\int_{-a}^{a} \log^+ |f(\rho e^{i\theta})|\, d\theta \leqq M$, $0 < \rho < 1$. *Then* $f(z)$ *is* B. T. *in the sector* S.[9]

Proof. Let S_n be a sequence of "truncated" sectors

$$S_n[1 - r_n \leqq \rho \leqq r_n < 1; -\alpha_n \leqq \theta = \arg z \leqq \alpha_n = \alpha - (1 - r_n)]$$

where, for $n = 1, 2, \cdots$, $0 < r_n \leqq r_{n+1} < 1$, $1 - r_n < \alpha$, and $\lim_{n\to\infty} r_n = 1$. The sequence S_n clearly exhaust S. We note that $d_n = 1 - r_n$ (see Theorem II.). Furthermore, conditions (a), (b) and (c) of Theorem II are satisfied where the point $a = 1$ (see footnote 7). Now

$$\iint_{S_n} \frac{\log^+ |f(z)|}{|1 - z|^2}\, dx dy = \int_{1-r_n}^{r_n} \int_{-\alpha_n}^{\alpha_n} \frac{\log^+ |f(\rho e^{i\theta})| \rho\, d\rho\, d\theta}{|1 - \rho e^{i\theta}|^2}$$

$$\leqq \int_{1-r_n}^{r_n} \frac{\rho}{(1-\rho)^2} \left[\int_{-\alpha_n}^{\alpha_n} \log^+ |f(\rho e^{i\theta})|\, d\theta\right] d\rho \leqq M \int_{1-r_n}^{r_n} \frac{\rho}{(1-\rho)^2}\, d\rho$$

the last inequality following by hypothesis. We now have

$$\iint_{S_n} \log^+ |f(z)|/|1 - z|^2\, dx\, dy \leqq M \int_0^{r_n} d\rho/(1-\rho)^2 \leqq M/(1 - r_n) = M/d_n.$$

Thus, condition (3. 1) of Theorem II is satisfied and hence $f(z)$ is B. T. in S.

5. Some criteria for the half-plane. In the present section we shall give two applications of Theorem I which are concerned with stating necessary and sufficient conditions for an analytic function to be B. T. in $x > 0$. As a first example we shall establish the following theorem.

THEOREM V. *Let* $f(z)$ *be analytic in* $x > 0$. *Suppose that for all* θ, $-\pi/2 < \theta < \pi/2$,

$$\int_{1/R}^{R} \log^+ |f(z)|/|1 + z|^2\, r\, dr \leqq MR, \qquad (z = re^{i\theta}). \tag{5. 1}$$

[9] This theorem is somewhat similar to a theorem due to I. Privaloff [7] for functions subharmonic in a sector, i. e. *If* $v(z)$ *is a subharmonic function in the open sector* $\alpha < \arg z < \beta$, $|z| < 1$ *and if* $\int_\alpha^\beta v^+(re^{i\theta})\, d\theta = 0(1)$, *then at almost any point of the arc* (α, β) *of the unit circle the limit of* $v(z)$ *exists along any path not touching the arc.* [$v^+(z) = v(z)$ *if* $v(z) \geqq 0$; $v^+(z) = 0$ *if* $v(z) < 0$].

Then $f(z)$ is B. T. *in $x > 0$ and $f(z) = b(z)g(z)$. Furthermore for all θ, $-\pi/2 < \theta < \pi/2$,*

$$\int_{1/R}^{R} \log^+ |g(z)|/|1+z|^2 \, r\,dr \leqq NR. \tag{5.2}$$

Conversely, if $f(z)$ is B. T. *in $x > 0$, then* (5.2) *holds with g replaced by f.*

Proof. We integrate (5.1) with respect to θ and find

$$\int_{-\pi/2}^{\pi/2} \int_{1/R}^{R} \log^+ |f(z)|/|1+z|^2 \, r\,dr\,d\theta \leqq \pi MR.$$

Consider the circles C, in the semi-circular ring A: $(-\pi/2 \leqq \theta \leqq \pi/2$, $1/R \leqq |z| \leqq R)$ which have their centers on the real axis and pass through the points $x = 1/R$ and $x = R$. The radii of these circles are $\bar{R} = \dfrac{(R+1/R)}{2}$. We see that $R \leqq 2\bar{R}$. Thus

$$\iint_C \log^+ |f(z)|/|1+z|^2 \, dx\,dy \leqq 2\pi M\bar{R}.$$

Since the circles C exhaust $x > 0$, the conditions of Theorem I are fulfilled. Hence $f(z)$ is B. T. in $x > 0$.

To establish the condition on $g(z)$ assuming $f(z)$ is B. T., we proceed as follows. From the representation for $\log |g(z)|$ given by (2.2) we see that

$$\log^+ |g(z)| \leqq \frac{1}{\pi} \int_{-\infty}^{\infty} [x^2 + (y-t)^2]^{-1} x \, |d\mu(t)| + |\lambda| \, x. \tag{5.3}$$

Multiplying both sides of (5.3) by $r/(1+r^2)$ and integrating with respect to r, we find

$$\int_{1/R}^{R} \frac{\log^+ |g(z)|}{1+r^2} \, r\,dr \tag{5.4}$$

$$\leqq \frac{1}{\pi} \int_{-\infty}^{\infty} \left[\int_{1/R}^{R} \frac{xr}{x^2+(y-t)^2} \cdot \frac{dr}{1+r^2} \right] |d\mu(t)| + |\lambda| \int_{1/R}^{R} \frac{xr\,dr}{1+r^2}.$$

Let us consider the integral in the brackets of (5.4), i. e.

$$I = \int_{1/R}^{R} \frac{xr}{x^2+(y-t)^2} \cdot \frac{dr}{1+r^2}. \tag{5.5}$$

To estimate this integral we consider two cases:

(a) $0 < |t| < 2R$. We employ contour integration to estimate I. It is seen that

$$I \leqq \mathrm{Re}\left[\int_{-\infty}^{\infty} \frac{r}{z-it}\,\frac{dr}{1+r^2}\right]. \tag{5.6}$$

Let us consider the complex integral $\int \frac{\zeta}{\zeta e^{i\theta}-it}\,\frac{d\zeta}{1+\zeta^2}$, the contour over which the integration takes place being specified below.

(i) For $t>0$ we consider a contour in $\mathrm{Im}(\zeta) \leqq 0$ defined by the semicircle C_R: $(|\zeta| = R)$ and the line segment $-R \leqq \zeta \leqq R$. We thus have

$$\int_{1/R}^{R} \frac{\zeta}{\zeta e^{i\theta}-it}\,\frac{d\zeta}{1+\zeta^2} + \int_{C_R} \frac{\zeta}{\zeta e^{i\theta}-it}\,\frac{d\zeta}{1+\zeta^2} = -2\pi i S$$

where S is the residue of the integrand of the complex integral at its only pole $\zeta = -i$ in $\mathrm{Im}(\zeta) \leqq 0$. Thus $\int_{-R}^{R} = \pi/(e^{i\theta}+t) - \int_{C_R}$. Letting $R \to \infty$ and noting that the integral over C_R vanishes, we have

$$\int_{-\infty}^{\infty} \frac{\zeta}{\zeta e^{i\theta}-it}\,\frac{d\zeta}{1+\zeta^2} = \frac{\pi}{e^{i\theta}+t}. \tag{5.7}$$

Taking the real part of both sides of (5.7) we have from (5.6) that

$$I \leqq \pi(\cos\theta + t)/(1+t^2+2t\cos\theta)$$

As $t>0$ and $-\pi/2 < \theta < \pi/2$, we see that $t\cos\theta \geqq 0$. Further as $|t| < 2R$, we find

$$I \leqq N_1 R/(1+t^2) \tag{5.8}$$

(ii) For $t<0$ we consider a contour in $\mathrm{Im}(\zeta) \geqq 0$ defined by the semicircle C_R: $(|\zeta| = R)$ and the line segment $-R \leqq \zeta \leqq R$. An argument similar to that given in (i) will show that

$$I \leqq N_2 R/(1+t^2) \tag{5.9}$$

(b) $|t| \geqq 2R > 1$. We return to (5.5). Using polar coordinates we find

$$I = \cos\theta \int_{1/R}^{R} \frac{r^2}{r^2+t^2-2rt\sin\theta}\,\frac{dr}{1+r^2}, \tag{5.10}$$

Now $(r^2 + t^2 - 2rt\sin\theta) \geqq (|t| - r)^2 \geqq t^2/4$. Since $1/t^2 \leqq 2/(1+t^2)$ we obtain on estimating (5. 10)

$$I \leqq 8R/(1+t^2). \tag{5.11}$$

Since the second integral of (5. 4) $\leqq R$, we have from (5. 8), (5. 9), and (5. 11) by splitting the integral of (5. 4) into the cases discussed that

$$\int_{1/R}^{R} \log^+ |g(z)|\, r/(1+r^2)dr \leqq N_3 R \int_{-\infty}^{\infty} |d\mu(t)|/(1+t^2) + |\lambda| R. \tag{5.12}$$

But the integral on the right side of (5. 12) is convergent. Hence the result is established. The converse of the theorem follows from this result.

As a second example we present a new and rather simple proof of a theorem due to A. Wishard [11], Theorem C, p. 672. This theorem was originally stated for a function meromorphic in a half plane. In this paper. however, we shall only establish this theorem for analytic functions.

THEOREM VI. *Let $f(z)$ be analytic in $x > 0$. Suppose that for all $x > 0$*

$$\int_{-\infty}^{\infty} \log^+ |f(z)|/|1+z|^2\, dy \leqq M, \qquad (z = x+iy). \tag{5.13}$$

Then $f(z)$ is B. T. *in $x > 0$; $f(z) = b(z)g(z)$ and for all $x > 0$*

$$\int_{-\infty}^{\infty} \log^+ |g(z)|/|1+z|^2\, dy \leqq N. \tag{5.14}$$

Conversely, if $f(z)$ is B. T. *in $x > 0$, then* (5. 14) *holds with g replaced by f.*

Proof. Let us consider the sequences of circles C_n, radii R_n, whose centers $a_n = R_n + \delta_n$ lie on the real-axis and $\delta_n > 0$. If $\delta_n \to 0$ and $R_n \to \infty$, this sequence will exhaust $x > 0$. Noting that the integral of

$$(\log^+ f(z))/|1+z|^2$$

taken over the area of the circle C_n is less than that over the circumscribed square, rewriting the integral over C_n as an iterated integral and employing (5. 13) we find

$$\int_{C_n} \log^+ |f(z)|/|1+z|^2\, dx\, dy \leqq 2MR_n.$$

Hence by Theorem I, it follows that $f(z)$ is B. T. in $x > 0$.

By means of (5. 3) we shall establish (5. 14). We find that

$$(5.15)\qquad \int_{-T_1}^{T_2} \frac{\log^+ |g(z)|}{|1+z|^2}\,dy \leqq \frac{1}{\pi}\int_{-T_1}^{T_2}\int_{-\infty}^{\infty} \frac{x\,|d\mu(t)|}{x^2+(y-t)^2}\,\frac{dy}{|1+z|^2} + |\lambda| \int_{-T_1}^{T_2} \frac{x\,dy}{|1+z|^2}.$$

Interchanging the order of integration in (5.15) and noting that

$$\frac{1}{1+x}\cdot\frac{2x+1}{(1+2x)^2+t^2} = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{1}{(1+x)^2+y^2}\cdot\frac{x\,dy}{x^2+(y-t)^2}$$

we find

$$(5.16)\qquad \int_{-T_1}^{T_2} \frac{\log^+ |g(z)|}{|1+z|^2}\,dy \leqq \frac{2x+1}{x+1}\int_{-\infty}^{\infty} \frac{|d\mu(t)|}{(1+2x)^2+t^2} + |\lambda| \int_{-\infty}^{\infty} \frac{x\,dy}{|1+z|^2}.$$

We note from Lemma I that the integral $\int_{-\infty}^{\infty} |d\mu(t)|/(1+t^2)$ is convergent. Furthermore $\int_{-\infty}^{\infty} dy/|1+z|^2 = \pi/(1+x)$. Thus, for all $x>0$ and any T_1 and T_2,

$$\int_{-T_1}^{T_2} \log^+ |g(z)|/|1+z|^2\,dy \leqq N.$$

Now let $T_1 \to \infty$ and $T_2 \to \infty$. Hence (5.14) is established. The converse follows easily from this result. For if $f(z)$ is B. T., then $|f| = |b|\,|g| < |g|$, but $\log^+ |g(z)|$ satisfies (5.3) and hence the conclusion follows.[10]

UNIVERSITY OF MINNESOTA.

---

[10] The O. N. Theorem states that if $F(w)$ is analytic in $|w|<1$ and

$$\int_0^{2\pi} \log^+ |F(re^{i\theta})|\,d\theta \leqq M$$

for all r, $0 \leqq r < 1$, then $F(w) = B(w)G(w)$ and

$$\int_0^{2\pi} \log^+ |G(re^{i\theta})|\,d\theta \leqq M$$

for $0 \leqq r < 1$ with *exactly the same constant* M. If Theorem VI is restated for the unit circle, it is seen that the vertical lines correspond to a sequence of circles passing through the point $w=1$ which exhaust the unit circle. From this point of view it would be natural to ask whether the constants M and N in (5.13) and (5.14) are the same also. An example can be given for which $M<N$ (see [1]).

BIBLIOGRAPHY

[1] A. Franck,, On analytic functions of bounded type, Doctoral Thesis, University of Minnesota, June 1949.

[2] R. M. Gabriel, "An inequality concerning the integrals of positive subharmonic functions along certain curves," *Journal of the London Mathematical Society*, vol. 4 (1929), pp. 133-139.

[3] W. Kryloff, "Uber Funktionen, die in der Halbebene regulär sind," *Mat. Sbornik*, new series volume 6 [old 48] (1939), pp. 95-138.

[4] F. and R. Nevanlinna, "Über die Eigenschaften analytischer Funktionen in der Umgebung einer singulären Stelle oder Linie," *Acta Soc. Sc. Fennicae*, vol. 50 (1922), No. 5.

[5] R. Nevanlinna, "Über eine Klasse meromorpher Funktionen," *Mathematische Annalen*, vol. 92 (1924), pp. 145-154.

[6] A. Ostrowski, "Über die Bedeutung der Jensenschen Formel fur einige Fragen der komplexen Funktiontheorie," *Acta Litt. ac Scient. Regiae Univ.* Hung. Francisco-Josephinae, vol. 1 (1923), pp. 80-87.

[7] I. Privaloff, "Sur un problème limite de la théorie des fonctions," *Bull. Acad. Sci. USSR*, vol. 7 (1935), pp. 301-304.

[8] W. Seidel, "Über die Ränderzuordnung bei konformen Abbildungen," *Mathematische Annalen*, vol. 104 (1931), pp. 182-242.

[9] C. Visser, "De conforme afbeelding van een speciale klasse van gebieden," *Handelingen van het XXVe Nederlandsch Natuur-en Geneeskundig Congress*, Delft, 1935.

[10] S. E. Warschawski, "Uber das Randverhalten der Ableitung der Abbilddungsfunktion bei konformer Abbildung," *Mathematische Zeitshrift*, vol. 35 (1932), pp. 322-456.

[11] A. Wishard, "Functions of bounded type," *Duke Mathematical Journal*, vol. 9 (1942), pp. 663-676.

INTEGRAL GEOMETRY IN HERMITIAN SPACES.*

By L. A. SANTALÓ.

1. Introduction. Let P_n be the n-dimensional complex projective space with the homogeneous coordinates x^i $(i = 0, 1, 2, \cdots, n)$. Let x^{*i} denote the complex conjugate of x^i.

We consider the group U (unitary group) of linear transformations

$$x'^i = \sum_{k=0}^{n} \xi_k{}^i x^k \tag{1.1}$$

which leaves the form

$$(xx^*) = \sum_{i=0}^{n} x^i x^{*i} \tag{1.2}$$

invariant. Here and in the sequel, a^* denotes the complex conjugate $(= \bar{a})$ of a. The coefficients of $\xi_k{}^i$ satisfy, then, the relations

$$(\xi_k \xi^*{}_l) = \sum_{i=0}^{n} \xi_k{}^i \xi^*{}_l{}^i = \delta_{kl}, \tag{1.3}$$

where $\delta_{kl} = 0$ if $k \neq l$ and $= 1$ if $k = l$.

Since the coefficients $\xi_k{}^i$ are complex numbers, the group U depends upon $(n+1)^2$ real parameters.

The geometry of P_n with the fundamental group U of transformations is called the Hermitian geometry (more precisely the elliptic hermitian geometry) and the space P_n itself is called an Hermitian space. The Integral Geometry in these spaces was initiated by Blaschke [2] who gave the densities for linear subspaces, without making applications to integral formulas. The case $n = 2$ was first considered by Varga [8] and later, in a more complete form, by Rohde [6].

In the present paper we generalize to the n-dimensional case some of the results which Varga and Rohde obtained for the plane. The main results we obtain are the following:

We first determine the explicit form of the left invariant element of volume du of the group U. If $L_r{}^0$ is a fixed linear subspace of dimension r (throughout the paper we shall mean by *dimension* the complex dimension)

* Received August 27, 1951.

and Γ_r is the subgroup of U which leaves L_r^0 invariant, we determine also the explicit form of the invariant element of volume dL_r of the homogeneous space U/Γ_r. Then, an easy calculation gives the total volumes (6.3) and (6.4) of U and U/Γ_r.

If the coordinates x^i are normalized such that $(xx^*) = 1$ it is well known that every variety C_p of dimension p $(p < n)$ has an invariant integral of degree $2p$, namely $J_p(C_p)$ defined in section **4**. Let C_h and C_r $(r + h \geq n)$ be two analytic varieties. Let uC_r denote the transform of C_r by the transformation u of U, and let $C_h \cap uC_r$ denote the $(h + r - n)$-dimensional variety intersection of C_h and uC_r. If the invariant element of volume du of U is normalized in such a way that the total volume of U is equal 1, we prove the following integral formula:

$$\int_U J_{h+r-n}(C_h \cap uC_r)\, du = J_h(C_h) J_r(C_r). \tag{1.4}$$

If instead of C_r we consider a linear subspace L_r, and the invariant elemert of volume dL_r in U/Γ_r is normalized such that the total volume of U/Γ_r is equal 1, we also have

$$\int_{U/\Gamma_r} J_{h+r-n}(C_h \cap L_r)\, dL_r = J_h(C_h). \tag{1.5}$$

We do not consider the case in which C_h is non-analytic. For $n = 2$ this more general case has been considered by Rohde [6], and it remains an interesting open question to extend the results of Rohde to the n-dimensional case.

2. Relative components and equations of structure of the unitary group. The coefficients $\xi_k{}^i$ of the unitary transformation (1.1) can be interpreted as coordinates of the transformed points ξ_k $(k = 0, 1, \cdots, n)$ of the $n + 1$ vertices of the n-simplex of reference; according to (1.3) they will be the vertices of an autoconjugate n-simplex with respect to the fundamental quadric $(\xi\xi^*) = 0$. These autoconjugate n-simplexes may be considered as the "frames" for the unitary group U (according to the theory of Cartan [3]) and the relative components ω_{jk} will then be defined by the equations

$$d\xi_k = \sum_{j=0}^{n} \omega_{kj} \xi_j. \tag{2.1}$$

From (2.1) and (1.3) de deduce

$$\omega_{kj} = (\xi^*{}_j d\xi_k), \tag{2.2}$$

and the equations of structure, deduced from (2.2) by remembering that $d(d\xi_k) = 0$, are

$$d\omega_{ih} = \sum_{k=0}^{n} [\omega_{ik}\omega_{kh}], \tag{2.3}$$

where here and throughout the paper square brackets denote exterior multiplication [3].

Notice that if ω^*_{kj} denotes the complex conjugate of the pfaffian form ω_{kj}, from (1.3) and (2.2) we deduce

$$\omega_{kj} + \omega^*_{jk} = 0. \tag{2.4}$$

3. Density for linear subspaces and cinematic density. We wish to define a measure for sets of r-dimensional linear subspaces L_r invariant with respect to U. If $L_r{}^0$ is a fixed L_r, and Γ_r denotes the subgroup of U which leaves $L_r{}^0$ invariant, then the problem is equivalent to the determination of an invariant element of volume in the homogeneous space U/Γ_r.

We follow the general method [5], [7]. Let $L_r{}^0$ be defined by the points $\xi_0, \xi_1, \cdots, \xi_r$. If it is fixed, in the equations (2.1) we will have

$$\omega_{jk} = 0 \quad \text{for} \quad 0 \leq j \leq r, \quad r+1 \leq k \leq n. \tag{3.1}$$

Since ω_{jk} are complex pfaffian forms, from $\omega_{jk} = 0$ we deduce $\omega^*_{jk} = 0$. Consequently the number of forms ω_{jk}, ω^*_{jk} $(0 \leq j \leq r,\ r+1 \leq k \leq n)$ is $2(r+1)(n-r)$. The density for sets of L_r, that is, the invariant element of volume in the space U/Γ_r will be, up to a constant factor,

$$dL_r = [\prod \omega_{jk}\omega^*_{jk}], \tag{3.2}$$

where the exterior product is taken always in absolute value and the indices range between the limits

$$0 \leq j \leq r, \qquad r+1 \leq k \leq n. \tag{3.3}$$

According to the equations of structure (2.3), it is easy to verify that $d(dL_r) = 0$, which is a sufficient condition in order that dL_r be effectively a density [7].

The invariant element of volume du in the space of the group U ("cinematic density" in the nomenclature of Blaschke) will be, up to a constant factor,

$$du = [\prod \omega_{jk}\omega^*_{jk} \prod \omega_{hh}] \tag{3.4}$$

where $j < k$, $0 \leq j, k, h \leq n$, that is, according to (2.4), du is equal to the absolute value of the exterior product of all the relative components.

4. The invariant integral $J_r(C_r)$ of an analytic variety C_r. Let C_r be an analytic variety of complex dimension r, that is, a variety defined by a set of $n+1$ analytic functions $x^i = x^i(t_1, t_2, t_3, \cdots, t_r)$ $(i = 0, 1, 2, \cdots, n)$ of r complex variables $t_1, t_2, \cdots, t_r$ in a domain D.

Assuming the homogeneous coordinates x^i normalized such that

$$(xx^*) = \sum_{i=0}^{n} x^i x^{*i} = 1, \tag{4.1}$$

let us consider the following differential form of degree $2r$:

$$\Omega^r = \sum [dx^{i_1} dx^{*i_1} \cdots dx^{i_r} dx^{*i_r}], \tag{4.2}$$

the summation being extended over all the combinations of $i_1, i_2, \cdots, i_r$ from 1 to n.

It is well known [4] that Ω^r is the only differential form of degree $2r$ which is invariant with respect to the group U.

The integral of Ω^r over an r-dimensional linear subspace L_r has the value [4],

$$\int_{L_r} \Omega^r = (2\pi i)^r / r!. \tag{4.3}$$

For a general analytic variety C_r we introduce the invariant integral $J_r(C_r)$ defined by

$$J_r(C_r) = r!/(2\pi i)^r \int_{C_r} \Omega^r. \tag{4.4}$$

If C_r is an algebraic variety, $J_r(C_r)$ coincides with its order (Cartan [4]).

For some purpose it is convenient to write Ω^r in another form. Let T_r be the r-dimensional linear subspace tangent to C_r at the point x, and take in T_r r points α_p $(p = 1, 2, \cdots, r)$ such that

$$(x\alpha^*_p) = 0, \qquad (\alpha_p \alpha^*_q) = \delta_{pq}. \tag{4.5}$$

We will have

$$dx^i = ax^i + \sum_{p=1}^{r} b_p \alpha_p{}^i, \qquad dx^{*i} = a^* x^{*i} + \sum_{p=1}^{r} b^*{}_p \alpha^*{}_p{}^i \tag{4.6}$$

where a and b_p are pfaffian forms given by

$$a = (x^* dx), \qquad b_p = (\alpha^*_p dx). \tag{4.7}$$

From (4.6) and (4.5) we deduce

$$\sum_{i=0}^{n} [dx^i dx^{*i}] = [aa^*] + \sum_{p=1}^{r} [b_p b^*{}_p]. \tag{4.8}$$

Since $a = -a^*$, we have $[aa^*] = 0$, and taking the r-th power of both sides of (4.8) we get $\Omega^r = [b_1 b^*{}_1 b_2 b^*{}_2 \cdots b_r b^*{}_r]$, or, according to (4.7),

$$\Omega^r = [\prod_{p=1}^{r} (\alpha_p^* dx)(\alpha_p dx^*)], \tag{4.9}$$

a formula which will be useful in the following sections.

5. The invariant integral J_r for the variety generated by the $(r-1)$-osculating spaces of a given analytic curve. As an example, we wish to evaluate J_r for the variety generated by the $(r-1)$-osculating linear spaces of a given analytic curve of the n-dimensional complex projective space.

Let C be an analytic curve defined by the $n+1$ parametric equations $y^i = y^i(t)$ $(i = 0, 1, 2, \cdots, n)$, where $y^i(t)$ are analytic functions of the complex variable t. In order to evaluate $J_1(C)$ we must normalize the coordinates y^i as indicated in (4.1). We set $x^i = y^i/(yy^*)^{\frac{1}{2}}$; then

$$dx^i = \{2(yy^*)dy^i - (dyy^*)y^i - (ydy^*)y^i\}/2(yy^*)^{3/2},$$
$$dx^{*i} = \{2(yy^*)dy^{*i} - (dy^*y)y^{*i} - (y^*dy)y^{*i}\}/2(yy^*)^{3/2},$$

and by exterior multiplication and addition,

$$\sum_{i=0}^{n} [dx^i dx^{*i}] = \{(yy^*)\Sigma [dy^i dy^{*i}] - [(y^*dy)(ydy^*)]\}/(yy^*)^2.$$

Notice that

$$(yy^*)\Sigma[dy^i dy^{*i}] - [(y^*dy)(ydy^*)]$$
$$= \{(yy^*)(y'y^{*\prime}) - (y^*y')(yy^{*\prime})\}[dtdt^*] = ((y \wedge y')(y^* \wedge y^{*\prime}))[dtdt^*],$$

where the notation $a \wedge b$ denotes the bivector with the components

$$\alpha^{ij} = a^i b^j - a^j b^i.$$

Consequently we have

$$\Omega^1 = \sum_{i=0}^{n} [dx^i dx^{*i}] = |\, y \wedge y' \,|^2 / |\, y \,|^4 [dtdt^*] \tag{5.1}$$

and finally, according to the definition (4.4),

$$J_1(C) = 1/2\pi i \int_C |\, y \wedge y' \,|^2 / |\, y \,|^4 [dtdt^*]. \tag{5.2}$$

If C is an algebraic curve, then (5.2) gives the order of C (except the sign which depends upon the orientation assumed for C).

Let us now consider the variety C_r generated by the $(r-1)$-osculating linear spaces of C. In order to evaluate $\int_{C_r} \Omega^r$ we consider two consecutive $(r-1)$-osculating spaces, $L_{r-1}(t)$, defined by the points $y(t), y'(t), \cdots, y^{(r-1)}(t)$, and $L_{r-1}(t+dt)$, defined by $y(t+dt), y'(t+dt), \cdots, y^{(r-1)}(t+dt)$, which are contained in the linear space $L_r(t)$ defined by the points $y(t), y'(t), \cdots, y^{(r)}(t)$. The angle $d\tau$ between $L_{r-1}(t)$ and $L_{r-1}(t+dt)$ is equal to the distance Ω^1 between the poles of $L_{r-1}(t)$ and $L_{r-1}(t+dt)$ considered as linear subspaces of L_r. Choosing the coordinate system so that the equations of L_r are $y^{r+1} = y^{r+2} = \cdots = y^n = 0$, the pole of $L_{r-1}(t)$ in $L_r(t)$ will be the point Y_r whose coordinates are the determinants of order r in the r by $r+1$ matrix $(y^{k(i)})$, where $i = 0, 1, \cdots, r-1$, and $k = 0, 1, \cdots, r$.

We now apply (5.1). The coordinates of the point Y'_r are the determinants of order r of the r by $r+1$ matrix $(y^{k(i)})$, where $i = 0, 1, \cdots, r-2, r$, and $k = 0, 1, \cdots, r$, and following a device due to Ahlfors and H. Weyl ([1], [9], p. 144), one can prove that

$$|Y_r \wedge Y'_r| = |Y_{r-1}| \cdot |Y_{r+1}|, \tag{5.2}$$

where Y_{r-1} and Y_{r+1} have a meaning analogous to that of Y_r, i. e. they are the multivectors defined by $(y, y', \cdots, y^{(r-2)})$ and $(y, y', \cdots, y^{(r)})$ respectively.

The device consists of choosing a coordinate system in which $y, y', \cdots, y^{(r)}$ have the coordinates

$$y(y^0, 0, \cdots, 0), y'(y^{0'}, y^{1'}, 0, \cdots, 0), \cdots, y^{(r)}(y^{0(r)}, y^{1(r)}, \cdots, y^{r(r)}).$$

In this system of coordinates the relation (5.2) becomes trivial.

Therefore we have

$$d\tau = |Y_{r-1}|^2 \cdot |Y_{r+1}|^2 / |Y_r|^4 [dtdt^*].$$

Since according to (4.3) $\int d\tau = \int \Omega^1 = 2\pi i$ we have the relation

$$d\tau / 2\pi i = \Omega^r / \int_{L_r} \Omega^r$$

and consequently we get

$$J_r(C_r) = r!/(2\pi i)^r \int_{C_r} \Omega^r = 1/2\pi i \int_C d\tau = 1/2\pi i \int_C |Y_{r-1}|^2 \cdot |Y_{r+1}|^2 / |Y_r|^4 [dtdt^*]. \tag{5.3}$$

This expression is due to Ahlfors and H. Weyl [1], [9]. If C is an algebraic curve (5.3) give the classes of different orders.

For instance, for a plane algebraic curve $y^0 = y^0(t)$, $y^1 = y^1(t)$, $y^2 = y^2(t)$, the class is given (except for the sign) by the integral

$$J_2(C) = 1/2\pi i \int_C |\, y \,|^2 \cdot |\, yy'y'' \,|^2 / |\, y \wedge y' \,|^4 [dtdt^*],$$

where $|\, yy'y'' \,|$ denotes the absolute value of the determinant formed by the components of y, y', y''.

6. Total volume of the unitary group and of the homogeneous spaces U/Γ_r. In order to calculate the volume $\int_U du$ of the unitary group U with the invariant element of volume du given by (3. 4), we put $\xi_0{}^k = \rho_k e^{i\theta_k}$ and consequently

$$\omega_{00} = (\xi^*{}_0 d\xi_0) = \sum_{k=0}^{n} (\rho_k d\rho_k + i\rho_k{}^2 d\theta_k).$$

Since $(\xi^*{}_0\xi_0) = \sum_0^n \rho_k{}^2 = 1$, we have $\sum_0^n \rho_k d\rho_k = 0$, and therefore the integral of ω_{00} over all possible values of the variables has the value

$$\int \omega_{00} = i \sum_{k=0}^{n} \int \rho_k{}^2 d\theta_k = 2\pi i.$$

The same proof gives

$$\int \omega_{jj} = 2\pi i. \tag{6. 1}$$

On the other hand we observe that, according to (4. 9), the exterior product

$$[\prod_{h=1}^{n-r} \omega_{r,r+h}\omega^*{}_{r,r+h}] = [\prod_{h=1}^{n-r} (\xi^*{}_{r+h} d\xi_r)\,(\xi_{r+h} d\xi^*{}_r)]$$

denotes the element Ω^{n-r} relative to the linear $(n-r)$-dimensional space defined by the points $\xi_r, \xi_{r+1}, \cdots, \xi_n$. Consequently (4. 3) gives

$$\int [\prod_{h=1}^{n-r} \omega_{r,r+h}\omega^*{}_{r,r+h}] = (2\pi i)^{n-r}/(n-r)\,!. \tag{6. 2}$$

From (3. 4), (6. 1) and (6. 2) we get

$$\int_U du = (2\pi i)^{n+1} \prod_{r=0}^{n} (2\pi i)^{n-r}/(n-r)\,!$$

which can be written

$$\int_U du = \prod_{h=1}^{n+1} (2\pi i)^h/(h-1)!. \tag{6.3}$$

This formula gives the total volume of the unitary group U.

We wish now to calculate the total volume of the homogeneous space U/Γ_r, or, what is the same, the total measure of all the linear r-dimensional spaces contained in the hermitian n-dimensional space.

Let us write, for the moment, $du = du_n$, in order to exhibit clearly the dimension n of the space. According to (3.2) and (3.4), we have the following relation between the elements of volume du_n, dL_r, and the elements du_r, du_{n-r}, of the r- and $(n-r)$-dimensional unitary groups:

$$du_n = [du_r\, du_{n-r-1}\, dL_r].$$

Integration of both sides of this equality over all possible values of the variables, taking (6.3) into account, gives

$$\prod_{h=1}^{n+1} (2\pi i)^h/(h-1)! = \prod_{h=1}^{r+1} (2\pi i)^h/(h-1)! \prod_{h=1}^{n-r} (2\pi i)^h/(h-1)! \int_{U/\Gamma_r} dL_r,$$

and consequently

$$\int_{U/\Gamma_r} dL_r = (2\pi i)^{(n-r)(r+1)} 1!2! \cdots r!/[n!(n-1)! \cdots (n-r)!]. \tag{6.4}$$

This is the measure of all the linear r-dimensional spaces of the hermitian n-dimensional space, i. e. the total volume of the homogeneous space U/Γ_r.

The finite values (6.3) and (6.4) induce a normalization of the elements of volume du and dL_r in such a way that the total volumes of the corresponding spaces are equal to 1. These normalized elements will be

$$du' = \prod_{h=1}^{n+1} (h-1)!/(2\pi i)^h du \tag{6.5}$$

$$dL_r' = (2\pi i)^{-(n-r)(r+1)} n!(n-1)! \cdots (n-r)!/1!2! \cdots r![dL_r]. \tag{6.6}$$

7. Linear subspaces which intersect an analytic variety. Let C_h be a fixed analytic variety of complex dimension h. Let L_r^0 be a fixed r-dimensional linear subspace and $L_r = uL_r^0$ be the transform of L_r^0 by $u \in U$. We assume $r + h - n \geq 0$, and let $C_h \cap L_r$ be the $(r+h-n)$-dimensional variety intersection of C_h with L_r. We wish to evaluate the integral

$$I = \int_{U/\Gamma_r} J_{h+r-n}(C_h \cap L_r)\, dL_r'. \tag{7.1}$$

Let ξ_0 be a point of $C_h \cap L_r$. In order to define L_r, we may take the points $\xi_0, \xi_1, \cdots, \xi_{h+r-n}$ on the linear $(h+r-n)$-space tangent to the intersection $C_h \cap L_r$ at ξ_0, and the points $\alpha_{h+r-n+1}, \cdots, \alpha_r$ such that

$$(7.2) \qquad (\xi_i \xi^*_j) = \delta_{ij}, \qquad (\xi^*_i \alpha_l) = 0, \qquad (\alpha^*_l \alpha_m) = \delta_{lm}.$$

Let $\gamma_{r+1}, \cdots, \gamma_n$ be $n-r$ points such that

$$(7.3) \qquad (\gamma^*_p \xi_i) = 0, \qquad (\gamma^*_p \alpha_l) = 0, \qquad (\gamma^*_p \gamma_q) = \delta_{pq},$$

and let $\beta_{h+r-n+1}, \cdots, \beta_h$ be $n-r$ points on the tangent space to C_h at ξ_0 such that

$$(7.4) \qquad (\beta^*_a \beta_b) = \delta_{ab}, \qquad (\beta^*_a \xi_i) = 0.$$

In (7.2), (7.3), and (7.4) we agree on the range of indices

$$0 \leq i, j \leq h+r-n, \quad h+r-n+1 \leq l, m \leq r, \qquad r+1 \leq p, q \leq n$$
$$h+r-n+1 \leq a, b \leq h.$$

According to (3.2) and (2.2) we have

$$(7.5) \qquad dL_r = [\prod (\gamma^*_p d\xi_i)(\gamma_p d\xi^*_i) \prod (\gamma^*_p d\alpha_l)(\gamma_p d\alpha^*_l)],$$

the product being extended over all γ, ξ, α. Since $d\xi_0$ is on the tangent h-space to C_h, we have

$$d\xi_0 = \sum_i A_i \xi_i + \sum_a B_a \beta_a,$$

where, according to (7.2) and (7.4), $A_i = (\xi^*_i d\xi_0)$, $B_a = (\beta^*_a d\xi_0)$.

Consequently

$$(\gamma^*_p d\xi_0) = \sum_a (\gamma^*_p \beta_a)(\beta^*_a d\xi_0), \qquad (\gamma_p d\xi^*_0) = \sum_a (\gamma_p \beta^*_a)(\beta_a d\xi^*_0),$$

and

$$(7.6) \qquad [\prod_p (\gamma^*_p d\xi_0)(\gamma_p d\xi^*_0)] = \|(\gamma^*_p \beta_a)\| \cdot \|(\gamma_p \beta^*_a)\| [\prod_a (\beta^*_a d\xi_0)(\beta_a d\xi^*_0)].$$

According to (4.9), the differential invariant Ω^h referred to C_h and the differential invariant Ω^{r+h-n} referred to $C_h \cap L_r$ may be written

$$(7.7) \qquad \Omega^h = [\prod_i (\xi^*_i d\xi_0)(\xi_i d\xi^*_0) \prod_a (\beta^*_a d\xi_0)(\beta_a d\xi^*_0)],$$

$$(7.8) \qquad \Omega^{h+r-n} = [\prod_i (\xi^*_i d\xi_0)(\xi_i d\xi^*_0)].$$

Consequently, from (7.5) and (7.6) we have

$$(7.9) \qquad [\Omega^{h+r-n} dL_r] = [\theta \Omega^h],$$

where θ is a differential form which does not depend upon ξ_0. According to (6.6) the same formula holds, up to a constant factor, for dL_r' instead of dL_r. The integral of θ over all possible values of the variables gives a constant value c which does not depend upon C_h. Consequently, integration of (7.9) gives $I = cJ_h(C_h)$. The value of c can be found by considering the case in which C_h is a linear subspace L_h. In this case we have $J_{h+r-n}(L_h \cap L_r) = J_{h+r-n}(L_{h+r-n}) = 1$, $J_h(L_h) = 1$, and $\int dL_r' = 1$. Consequently $c = 1$.

We get the final result

$$\int_{U/T_r} J_{r+h-n}(C_h \cap L_r)\, dL_r' = J_h(C_h). \tag{7.10}$$

If $r + h - n = 0$, $J_0(C_h \cap L_r)$ denotes the number of intersection points of L_r with C_h.

8. Analytic varieties which intersect each other. Let C_h, C_r $(h + r - n \geqq 0)$ be two analytic varieties of dimension h, r respectively. Let uC_r be the transform of C_r by $u \in U$. We consider the integral

$$I = \int_U J_{r+h-n}(C_h \cap uC_r)\, du'. \tag{8.1}$$

The frame which determines u may be chosen in the following way. Let ξ_0 be a point of the intersection $C_h \cap uC_r$. We choose the points $\xi_1, \xi_2, \cdots, \xi_{h+r-n}$ on the linear $(h + r - n)$-dimensional space tangent to the variety $C_h \cap uC_r$ at ξ_0, and the points $\alpha_{h+r-n+1}, \cdots, \alpha_r$ on the tangent r-space to C_r at ξ_0, in such a way that

$$(\xi^*_i \xi_j) = \delta_{ij}, \qquad (\xi^*_i \alpha_l) = 0, \qquad (\alpha^*_l \alpha_m) = \delta_{lm}. \tag{8.2}$$

Let $\gamma_{r+1}, \cdots, \gamma_n$ be $n - r$ points such that

$$(\gamma^*_p \xi_i) = 0, \qquad (\gamma^*_p \alpha_l) = 0, \qquad (\gamma^*_p \gamma_q) = \delta_{pq}. \tag{8.3}$$

In (8.2), (8.3), and in the remainder of this section we agree on the ranges of indices $0 \leq i, j \leq h + r - n$, $h + r - n + 1 \leq l, m \leq r$, $r + 1 \leq p, q \leq n$.

The frame which determines u is the n-simplex ξ_i, α_l, γ_p. According to (3.4) we have

$$du = [\Pi(\xi^*_i d\xi_i)\Pi(\alpha^*_l d\alpha_l)\Pi(\gamma^*_p d\gamma_p)\Pi(\xi^*_i d\xi_j)(\xi_i d\xi^*_j)\Pi(\alpha^*_l d\alpha_m)(\alpha_l d\alpha^*_m)$$
$$\Pi(\alpha^*_l d\xi_i)(\alpha_l d\xi^*_i)\Pi(\gamma^*_p d\gamma_q)(\gamma_p d\gamma^*_q)\Pi(\gamma^*_p d\xi_i)(\gamma_p d\xi^*_i)\Pi(\gamma^*_p d\alpha_l)(\gamma_p d\alpha^*_l)], \tag{8.4}$$

where $i \neq j$, $l \neq m$, $p \neq q$.

Let $\beta_{h+r-n+1}, \cdots, \beta_h$ be $n-r$ points on the tangent h-space to C_h at ξ_0, such that

$$(8.5) \qquad (\beta^*_a \beta_b) = \delta_{ab}, \qquad (\beta^*_a \xi_i) = 0,$$

where $h + r - n + 1 \leq a, b \leq h$.

Since we always take ξ_0 on C_h, we have

$$d\xi_0 = \sum_i A_i \xi_i + \sum_a B_a \beta_a$$

where, according to (8.2) and (8.3), $A_i = (\xi^*_i d\xi_0)$, $B_a = (\beta^*_a d\xi_0)$.

Consequently we have

$$(\gamma^*_p d\xi_0) = \sum_a (\gamma^*_p \beta_a)(\beta^*_a d\xi_0), \qquad (\gamma_p d\xi^*_0) = \sum_a (\gamma_p \beta^*_a)(\beta_a d\xi^*_0),$$

and by exterior multiplication,

$$(8.6) \qquad [\prod_p (\gamma^*_p d\xi_0)(\gamma_p d\xi^*_0)] = \|(\gamma^*_p \beta_a)\| \cdot \|(\gamma_p \beta^*_a)\| [\prod_a (\beta^*_a d\xi_0)(\beta_a d\xi^*_0)].$$

The differential invariant Ω^r referred to C_r, according to (4.9), is

$$(8.7) \qquad \Omega^r = [\prod_i (\xi^*_i d\xi_0)(\xi_i d\xi^*_0) \prod_l (\alpha^*_l d\xi_0)(\alpha_l d\xi^*_0)].$$

From (7.7), (7.8), (8.4), (8.6), and (8.7) we deduce

$$(8.8) \qquad [\Omega^{h+r-n} du] = [\psi \Omega^r \Omega^h],$$

where ψ is a differential form which does not depend upon ξ_0. Up to a constant factor, according to (6.5), the same formula holds for du' instead of du. The integral of ψ over all possible values of the variables gives a constant value c. Consequently from (8.1) and (8.8) we deduce $I = cJ_h(C_h)J_r(C_r)$.

In order to determine the value of the constant c we consider the case in which C_h and C_r are linear spaces L_h and L_r. In this case we have $J_{h+r-n}(L_h \cap uL_r) = 1$, $\int du' = 1$, $J_h(C_h) = J_r(C_r) = 1$. Consequently we have proved the integral formula

$$(8.9) \qquad \int_U J_{r+h-n}(C_h \cap uC_r)\, du' = J_h(C_h) J_r(C_r).$$

If $r + h - n = 0$, $J_0(C_h \cap uC_r)$ denotes the number of intersection points of C_h with uC_r.

For algebraic varieties, since the invariant J coincides with the order of the variety, (8.9) is an integrated form of the theorem of Bezout. Therefore (8.9) may be considered as a generalization of the theorem of Bezout to analytic varieties.

If we have $p+1$ analytic varieties $C_{h_0}, C_{h_1}, \cdots, C_{h_p}$ such that $h_0 + h_1 + \cdots + h_p \geq np$, from (8. 9) we get immediately by recurrence

$$\int_U J_{h_0+h_1+\ldots+h_p-np}(C_h \cap u_1 C_{h_1} \cap u_2 C_{h_2} \cap \cdots \cap u_p C_{h_p})\, du_1' du_2' \cdots du_p' = J_{h_0}(C_{h_0}) J_{h_1}(C_{h_1}) \cdots J_{h_p}(C_{h_p}).$$

THE UNIVERSITY OF LA PLATA, ARGENTINA.

BIBLIOGRAPHY.

[1] L. V. Ahlfors, "The theory of meromorphic curves," *Acta Societatis Scientiarum Fennicae*, A, vol. III, 1941.

[2] W. Blaschke, "Densita negli spazi di Hermite," *Rendiconti dell' Academia dei Lincei*, vol. 29, 6, pp. 105-108, 1939.

[3] E. Cartan, *La théorie des groupes finis et continus et la géométrie differentielle traités par la méthode du repere mobile*, Paris, 1937.

[4] E. Cartan, "Sur les invariants intégraux de certains espaces homogènes clos et les proprietés topologiques de ces espaces," *Annales de la Société Polonaise des Mathématiques*, vol. 8, pp. 181-225, 1929.

[5] S. S. Chern, "On integral geometry in Klein spaces," *Annals of Mathematics*, vol. 43, pp. 178-189, 1942.

[6] H. Rohde, "Unitäre Integralgeometrie," *Hamburger Abhandlungen*, vol. 13, pp. 295-318, 1940.

[7] L. A. Santaló, "Integral Geometry in projective and affine spaces," *Annals of Mathematics*, vol. 51, pp. 739-755, 1950.

[8] O. Varga, "Ueber die Integralinvarianten die zu einer Kurve in der Hermitischen Geometrie gehören," *Acta Litt. Scientarum Szeged*, vol. 9, pp. 88-102, 1939.

[9] H. Weyl, "Meromorphic functions and analytic curves," *Annals of Mathematics Studies*, No. 12, Princeton, 1943.

ON THE PARABOLIC GENERATORS OF THE PRINCIPAL CONGRUENCE SUBGROUPS OF THE MODULAR GROUP.*

By EMIL GROSSWALD.

1. The set of matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, where the integers a, b, c, d satisfy

$$ad - bc = 1 \tag{1}$$

and where we consider $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -a & -b \\ -c & -d \end{pmatrix}$, form the modular group. We define $\Gamma(p)$, the principal congruence subgroup of Γ modulo a prime p, by the additional conditions $a \equiv d \equiv 1$ and $b \equiv c \equiv 0$ (unless otherwise stated, all congruences are taken mod p). For $p > 3$, H. Frasch has shown ([3]; see also [7]) that $\Gamma(p)$ can be generated as a free group by the generators $S^p = \begin{pmatrix} 1 & p \\ 0 & 1 \end{pmatrix}$ and $(p-1)p(p+1)/12$ other generators, depending on three parameters, identified by the symbols (λ, μ, ν) and defined as follows. Let

$$V = TS^{\alpha}TS^{\beta}TS^{\alpha} = \begin{pmatrix} \beta & \alpha\beta - 1 \\ -(\alpha\beta - 1) & -\alpha(\alpha\beta - 2) \end{pmatrix}, \tag{2}$$

where α and β are any two primitive roots (mod p) satisfying $\alpha\beta \equiv 1$. Let λ, μ, ν be any set of integers satisfying $0 \leqq \lambda \leqq (p-3)/2$, $0 \leqq \mu \leqq p-1$, $1 \leqq \nu \leqq p-1$, and define

$$\begin{gathered} \sigma = \operatorname*{ind}_{\alpha} \nu, \qquad \lambda_* = \lambda + \sigma \quad (\text{reduced mod } \tfrac{1}{2}(p-1)), \\ \mu_* = \mu\nu^2 - \nu \quad (\text{reduced mod } p), \quad \nu\nu_* \equiv -1 \pmod{p}; \end{gathered} \tag{3}$$

so that λ_*, μ_*, ν_* are in the same range, respectively, as λ, μ, ν. Let us call "reciprocals" two numbers related as ν and ν_*. Finally, let

$$\begin{aligned} R &= S^{\mu}TS^{\nu}TS^{-\nu_*}TS^{-\mu_*} \\ &= \begin{pmatrix} -\mu(\nu\nu_* + 1) + \nu_* & \mu\mu_*(\nu\nu_* + 1) - (\mu\nu + \mu_*\nu_*) + 1 \\ -(\nu\nu_* + 1) & \mu_*(\nu\nu_* + 1) - \nu \end{pmatrix}. \end{aligned} \tag{4}$$

Then the generators (λ, μ, ν) of $\Gamma(p)$ are defined by

* Received January 12, 1951.

$$(5) \qquad (\lambda, \mu, \nu) = V^{\lambda} R V^{-\lambda_\bullet}.$$

The object of this paper is to show that if the prime p satisfies a certain condition [1] to be stated later, the independent generators (λ, μ, ν) can be chosen in such a way that none of them is parabolic,[2] so that the only parabolic generator of $\Gamma(p)$ is S^p.

For our proof, we have to change slightly the process of elimination indicated by Frasch [3], whose method we follow, however, closely. Throughout the paper we consider only values of $p > 7$ (cf. a similar situation in [7], section I); in the other cases we would need sometimes special considerations, which complicate needlessly the exposition, as the corresponding generators are explicitly indicated in [3] and none of them is parabolic. A few inequalities will be meaningful only for some higher values of p and will be understood to apply only for them.

2. We recall some results of Frasch [3], with some of their immediate consequences. The $\frac{1}{2}p(p-1)^2$ generators (λ, μ, ν) satisfy $p(p-1)(5p-7)/12$ distinct defining relations, each permitting the elimination of one generator. We use them in order to eliminate all generators with a same value ν_0 of ν, unless ν_0 satisfies

$$(6) \qquad \nu_0^2 \equiv -1 \quad \text{or} \qquad (6') \qquad \nu_0(\nu_0 - 1) \equiv -1.$$

The defininig relations can be arranged in groups of binary and ternary relations. Following Frasch, we use the first group of binary relations to eliminate the generators with $\nu = 1$. The following ternary relations permit the elimination of $\nu = p-1$. Using the second group of binary relations, $\nu = 2$ and $\nu = p-2$ are eliminated; and the corresponding ternary relations permit the elimination of their reciprocals, $\nu = \frac{1}{2}(p \pm 1)$. In general, for $k \geqq 3$, we use the binary group to eliminate $\nu = k$, $\nu = p-k$, and $\nu = \frac{1}{2}\{p \pm (2k-3)\}$, and the corresponding ternary relations to eliminate their reciprocals. The generators with the mentioned values of ν actually appear in the defining relations in the stated order, as explicitly shown in [3]. The number of relations at our disposal is larger than the number of generators on which we have centered our attention, and we may use the remaining ones in the order suggested in [3]. The process of elimination

---

[1] It seems likely that all primes satisfy that condition; see last footnote.

[2] A corresponding property holds for the generators V_k of $\Gamma_0(p)$, defined in [9]. See [5].

has to be modified for $p \not\equiv -1 \pmod{12}$, when ν takes on values satisfying (6) or (6′). Let $\nu = k$ or $\nu = p - k$ satisfy (6). Then $k^2 + 1 \equiv 0$, so that $k^2 - (p-1) \geqq 0$, and $k \geqq (p-1)^{\frac{1}{2}}$. If ν satisfies (6′), then $k(k-1) \equiv -1$, so that $k^2 - k - (p-1) \geqq 0$, $k \geqq \frac{1}{2} + (p - \frac{3}{4})^{\frac{1}{2}} > (p-1)^{\frac{1}{2}}$. In the same way we show that if $\nu = \frac{1}{2}\{p \pm (2k-3)\}$ satisfies (6) or (6′), then $k > \frac{1}{2}(p^{\frac{1}{2}} + 1)$. It follows that all values of ν belonging to any one of the three intervals

$$(7) \qquad 1 \leqq \nu \leqq p^{\frac{1}{2}} - 1; \quad \tfrac{1}{2}(p - p^{\frac{1}{2}}) + 1 < \nu < \tfrac{1}{2}(p + p^{\frac{1}{2}}) - 1; \quad p - p^{\frac{1}{2}} + 1 < \nu \leqq p - 1$$

are not affected by the difficulties arising from (6), (6′), and can be eliminated from the final set of independent generators, using only the binary relations. Whenever we reach any value satisfying (6), or (6′), we continue as in [3]. If ν does not satisfy (6) or (6′), respectively, then the same is true of its reciprocal, and the generators with values of ν reciprocal to those of (7) will also be eliminated, using the corresponding ternary relations. The values of ν in (7) and their reciprocals form a set of at most $3(p^{\frac{1}{2}} - 1)$ distinct numbers. For the elimination of the corresponding generators we dispose (for some fixed λ and μ) of $\frac{1}{12}p(p-1)(5p-7)/\frac{1}{2}p(p-1) = (5p-7)/6$ relations. For $p > 7$, $(5p-7)/6 \geqq 3(p^{\frac{1}{2}} - 1)$, proving the previous assertion that the number of independent defining relations exceeds the number of generators which we want to eliminate. Finally, we remark that there remain in the generators only values of ν, which satisfy, with their reciprocals, the inequalities

$$(8) \qquad p^{\frac{1}{2}} - 1 < \nu, \nu_* < \tfrac{1}{2}(p - p^{\frac{1}{2}}) + 1, \text{ or } \tfrac{1}{2}(p + p^{\frac{1}{2}}) - 1 < \nu, \nu_* < p - p^{\frac{1}{2}} + 1.$$

3. From (3) and (5) it follows that

$$(9) \qquad \begin{aligned} &V^{-\lambda}(\lambda, \mu, \nu)\, V^{\lambda} = RV^{-\sigma}, \quad \text{if} \quad \lambda + \sigma < \tfrac{1}{2}(p-1), \quad \text{and} \\ &V^{-\lambda}(\lambda, \mu, \nu)\, V^{\lambda} = RV^{-\sigma + \frac{1}{2}(p-1)}, \quad \text{if} \quad \lambda + \sigma \geqq \tfrac{1}{2}(p-1). \end{aligned}$$

If (λ, μ, ν) is parabolic, then all its transforms $M^{-1}(\lambda, \mu, \nu)M$ by any modular matrix M are also parabolic. In order to prove that (λ, μ, ν) is not parabolic it is, therefore, sufficient to show that $RV^{-\sigma}$ and $RV^{-\sigma + \frac{1}{2}(p-1)}$ are not parabolic. Let, generally, $V^n = \begin{pmatrix} A_n & B_n \\ C_n & D_n \end{pmatrix}$. Then it follows from $V^{n+1} = V^n \cdot V = \begin{pmatrix} A_{n+1} & B_{n+1} \\ C_{n+1} & D_{n+1} \end{pmatrix}$ that

$$A_{n+1} = \beta A_n - (\alpha\beta - 1)B_n \qquad B_{n+1} = (\alpha\beta - 1)A_n - \alpha(\alpha\beta - 2)B_n$$
(10)
$$C_{n+1} = \beta C_n - (\alpha\beta - 1)D_n \qquad D_{n+1} = (\alpha\beta - 1)C_n - \alpha(\alpha\beta - 2)D_n.$$

For $n = 1$, by (2), $A_1 - D_1 = B_1(\alpha - (\alpha - \beta)/(\alpha\beta - 1))$ and $B_1 + C_1 = 0$. Assuming that

(11) $\qquad A_n - D_n = 2tB_n$, where $t = \frac{1}{2}(\alpha - (\alpha - \beta)/(\alpha\beta - 1))$,

(12) $\qquad B_n + C_n = 0$

hold for n, it follows by induction on n, using (10), that (11) and (12) hold for $n + 1$; they hold, therefore, for all n. It follows from (1) and (12) that

(13) $$A_nD_n + B_n^2 = 1$$

and, as B_n is an integer,

(14) $$A_nD_n \leqq 0.$$

Hence, by (11) and (12),

(15) $$A_nB_n > 0, \qquad A_nC_n < 0.$$

For $n = 1$, by (2), $B_1 > A_1$; assuming that, for arbitrary integral n,

(16) $$|B_n| > |A_n|,$$

it follows by induction on n, using (10) and (15), that (16) holds for $n + 1$ and so for all n. From (16) it obviously follows that $(\alpha\beta - 1)B_n/A_n - \beta > 0$ so that, by (10), $A_{n+1}A_n < 0$. Hence, using (15), (12) and (14), $B_nB_{n+1} < 0$, $C_nC_{n+1} < 0$ and $D_nD_{n+1} \leqq 0$. We shall see immediately that the equality cannot hold, so that we have, generally, with M standing for A, B, C or D,

(17) $$M_nM_{n+1} < 0.$$

Let, generally, $\alpha\beta = Kp + 1$, $K \geqq 1$; then $B_1 = Kp$. From (10) it follows, by (14) and (17), that $|B_{n+1}/B_n| = \alpha(\alpha\beta - 2) - (\alpha\beta - 1)A_n/B_n > \alpha(\alpha\beta - 2) - (\alpha\beta - 1) = \alpha(Kp - 1) - Kp \geqq Kp - 2$ so that $|B_n| > |B_1|(Kp - 2)^{n-1} = K^n p^n (1 - 2/Kp)^{n-1}$ and, for $p > 7$,

(18) $$|B_n| > (5p/7)^n.$$

Hence it follows in particular that $|B_n| > 1$ and the equality signs in (14) and in $D_nD_{n+1} \leqq 0$ may be suppressed. We have, consequently, using (15) and (12),

$$A_nD_n < 0, \qquad B_nD_n < 0, \qquad C_nD_n > 0. \tag{19}$$

We solve now (11) and (12) for A_n and D_n, using (16) and (19) to determine the signs of the radicals and obtain

$$A_n = B_n(t - (t^2 - 1 + B_n^{-2})^{\frac{1}{2}})$$

$$D_n = -B_n(t + (t^2 - 1 + B_n^{-2})^{\frac{1}{2}}).$$

As, by (11) and (18), $2 < 2t < B_n$, it follows that

$$t - 1/2t > (t^2 - 1 + B_n^{-2})^{\frac{1}{2}} > t - 1/(2t-1),$$

so that

$$\begin{gathered} 2t - 1/(2t-1) < |D_n/B_n| < 2t - 1/2t \\ 1/2t < A_n/B_n < 1/(2t-1). \end{gathered} \tag{20}$$

4. Let now, for integral $n \geqq 1$, $RV^{-n} = R\begin{pmatrix} D_n & -B_n \\ -C_n & A_n \end{pmatrix} = \begin{pmatrix} \mathfrak{A}_n & \mathfrak{B}_n \\ \mathfrak{C}_n & \mathfrak{D}_n \end{pmatrix}$.

If this matrix is parabolic, then we have (see [2, p. 21]) $|\mathfrak{A}_n + \mathfrak{D}_n| = 2$, where, by (4),

$$\begin{aligned} \mathfrak{A}_n &= -D_n[\mu(\nu\nu_* + 1) - \nu_*] - C_n[\mu\mu_*(\nu\nu_* + 1) - (\mu\nu + \mu_*\nu_*) + 1] \\ \mathfrak{D}_n &= B_n(\nu\nu_* + 1) + A_n[\mu_*(\nu\nu_* + 1) - \nu]. \end{aligned} \tag{21}$$

If $\mu\mu_* \neq 0$, all brackets are positive, as is obvious for all except the coefficient of C_n. Let its value be N, and apply (1) to R. It follows that $N > 0$, and by (15), (18) and (19), we obtain $|\mathfrak{A}_n + \mathfrak{D}_n| > |B_n|(N + \nu\nu_* + 1) > (5p/7)^n(N + p) > 2$, so that RV^{-n} cannot be parabolic. If $\mu = 0$, then, by (3), $\mu_* = p - \nu$. From (15) and (19) it follows that $|\mathfrak{A}_n + \mathfrak{D}_n| = ||\mathfrak{A}_n| - |\mathfrak{D}_n||$; and by (20) and (21), we have $l < |\mathfrak{A}_n/B_n| < L$, $H > |\mathfrak{D}_n/B_n| > h$, with

$$\begin{aligned} l &= (p - \nu)\nu_* - 1 + [2t - 1/(2t-1)]\nu_*, \\ L &= (p - \nu)\nu_* - 1 + (2t - 1/2t)\nu_*, \end{aligned}$$

$$h = (\nu\nu_* + 1) + [(p - \nu)(\nu\nu_* + 1) - \nu]/2t,$$

$$H = (\nu\nu_* + 1) + [(p - \nu)(\nu\nu_* + 1) - \nu]/(2t - 1).$$

Consequently,

$$|\mathfrak{D}_n| - |\mathfrak{A}_n| > |B_n|(h - L) \tag{22}$$

and we assert that

$$h - L > 1/2t \tag{23}$$

if

$$\text{either } \beta < \alpha \leqq \nu \text{ or } \alpha \leqq \nu - 1, \beta > \alpha. \tag{24}$$

Proof. (23) is equivalent to

$$4t^2 - 2t[2(\nu + 1/\nu_*) - p] - [(p - \nu)(\nu + 1/\nu_*) - (\nu + 1)/\nu_* + 1] \leqq 0,$$

which is satisfied for $0 \leqq t \leqq \frac{1}{2}\nu$, or, by (11), for $\alpha \leqq \nu + (\alpha - \beta)/(\alpha\beta - 1)$, proving the assertion.

By virtue of (8), (24) is always satisfied for

$$\alpha \leqq p^{\frac{1}{2}} - 2 \tag{25}$$

and also for

$$\alpha \leqq \tfrac{1}{2}(p + p^{\frac{1}{2}}) - 2, \text{ if } \tfrac{1}{2}(p + p^{\frac{1}{2}}) - 1 < \nu < p - p^{\frac{1}{2}} + 1. \tag{25'}$$

Similarly,

$$|\mathfrak{A}_n| - |\mathfrak{D}_n| > |B_n|(l - H) \tag{26}$$

where

$$l - H \geqq 1/(2t - 1) \text{ if } \qquad \alpha \geqq \nu + 2. \tag{27, 28}$$

Proof. (27) is equivalent to

$$4t^2 - 2t[2(\nu + 1/\nu_*) - p + 1] - [(p - \nu - 1)(\nu + 1/\nu_*) - (\nu - 2)/\nu_* + p - \nu + 1] \geqq 0,$$

which is satisfied for $2t \geqq \nu + 1/\nu_* + 1$, or, by (11), for $\alpha \geqq \nu + 1/\nu_* + 1 - (\alpha - \beta)/(\alpha\beta - 1)$ proving the assertion. In view of (8), α satisfies (28), provided that

$$\alpha \geqq p - p^{\frac{1}{2}} + 3 \tag{29}$$

and also for

$$(29') \qquad \alpha \geqq \tfrac{1}{2}(p - p^{\frac{1}{2}}) + 3, \text{ if } p^{\frac{1}{2}} - 1 < \nu < \tfrac{1}{2}(p - p^{\frac{1}{2}}) + 1.$$

Combining (25′) and (29′), it follows that for any possible value of ν either (23) or (27) holds, in case we have simultaneously [3]

$$(30) \qquad \tfrac{1}{2}(p - p^{\frac{1}{2}}) + 3 < \alpha < \tfrac{1}{2}(p + p^{\frac{1}{2}}) - 2.$$

We shall say that a prime p has the property C, in case [4] at least one of the following three alternatives holds:

(a) the smallest primitive root of p does not exceed $p^{\frac{1}{2}} - 2$;

(b) the largest primitive root of p is not less than $p - p^{\frac{1}{2}} + 3$;

(c) there exists a primitive root of p in the interval (30).

Let us choose for α a primitive root satisfying either (a), or (b), or (c). From (22), (23), (26), and (27) it follows now, on account of (18), that if p has the property C, then

$$(31) \qquad |\mathcal{A}_n + \mathcal{D}_n| > |B_n|/2t > (5p/7)^n/p.$$

By (3) $\alpha^\sigma \equiv \nu$ and, by the choice of α, $\nu \neq \alpha$; consequently, $\sigma \neq 1$ in (9), and it is sufficient to consider in (31) only values $n \geqq 2$. As for $n \geqq 2$, $(5p/7)^n/p > 2$, it follows from (31) that $|\mathcal{A}_n + \mathcal{D}_n| > 2$, proving that if $\mu = 0$, then RV^{-n} is not parabolic. The case $\mu_* = 0$ can be reduced to the previous one by interchanging ν and ν_*, which is possible without any further change of the argument or of the chosen value of α, since ν and ν_* were both eliminated from the intervals (7). This finishes the proof that RV^{-n} is not parabolic.

Similarly, let $RV^n = \begin{pmatrix} \mathcal{A}'_n & \mathcal{B}'_n \\ \mathcal{C}'_n & \mathcal{D}'_n \end{pmatrix}$. Assuming that this matrix is parabolic, we obtain, proceeding as before,

---

[3] (25), (29) and (30) may be improved to $\alpha < p^{\frac{1}{2}}$, $p - \alpha \leqq p^{\frac{1}{2}}$ and $|\alpha - \frac{1}{2}p| \leqq \frac{1}{2}p^{\frac{1}{2}}$ respectively, by sharper estimates in (20). We remark also that (30) is meaningful only for $p \geqq 29$. For all smaller primes, however, either (25), or (29) hold. See also remark at the end of 1.

[4] Although it seems likely, that every prime has the property C, the best that is known at present is a weaker result of Loo-keng Hua [6]; see also [1], [4], and [8]. In [7] J. Nielsen considers only primes p, admitting as primitive roots either 2, or $p-2$ (consequently also $\frac{1}{2}(p \pm 1)$). Property C represents a weakened Nielsen condition and every prime, satisfying Nielsen's condition, obviously has the property C.

$$
\begin{aligned}
|\,\mathcal{A}'_n + \mathcal{D}'_n\,| = &- A_n[\mu(\nu\nu_* + 1) - \nu_*] \\
(21') \qquad &+ C_n[\mu\mu_*(\nu\nu_* + 1) - (\mu\nu + \mu_*\nu_*) + 1] \\
&- B_n(\nu\nu_* + 1) + D_n[\mu_*(\nu\nu_* + 1) - \nu] = 2.
\end{aligned}
$$

We observe that, except for its sign, (21′) is obtained from the values (21), interchanging the couples (μ, ν) and (μ_*, ν_*). Consequently, $|\,\mathcal{A}_n + \mathcal{D}_n\,| \neq 2$ implies $|\,\mathcal{A}'_n + \mathcal{D}'_n\,| \neq 2$ contrary to (21′), and RV^n is not parabolic, unless $n = 1$. From $\frac{1}{2}(p+1) - \sigma = 1$ would follow, however, by (3), $\sigma = 0$ and, further, $\nu = \alpha^\circ = 1$, contradicting (7). Therefore, in (9), the exponent of V is different from one and $RV^{\frac{1}{2}(p+1)-\sigma}$ is not parabolic. By (9) we conclude that no generator (λ, μ, ν) is parabolic, so that the only parabolic generator of $\Gamma(p)$ is S^p, q. e. d.

Remark. A separate consideration of the cases $p \equiv \pm 1, \pm 5 \pmod{12}$ yields some additional information and suggests that the above result holds for all primes, whether they have property C or not, with at most a finite number of exceptions. As an example, let $p \equiv -1 \pmod{12}$. Then no relations (6), (6′) exist, and all values of ν of the form $\nu = k$, $\nu = p - k$, $1 \leqq k \leqq [(5p-7)/24]$, and their reciprocals can be eliminated from the generators. Erdös has shown [1] that every prime p has a primitive root $\alpha < (\log p)^{17} p^{\frac{1}{4}}$; and the considerations of **4** hold, provided $\alpha \leqq \nu - 1$, or

$$
(32) \qquad (\log p)^{17} p^{\frac{1}{4}} \leqq [(5p-7)/24] - 1,
$$

which is true for all sufficiently large p. In the same way, results of Fueter [4] could be used to improve our results in the cases $p \equiv +1$, and $p \equiv \pm 5 \pmod{12}$. The interest of such considerations is, however, considerably diminshed by the fact that "sufficiently large" means, for (32), $p > e^{200}$. Therefore, the additional results obtained do not seem to justify a detailed case distinction (mod 12). It is rather to be expected that the known results concerning the smallest primitive root (mod p) will be improved sufficiently (see [8]) to show that all primes have the property C.

UNIVERSITY OF SASKATCHEWAN.

BIBLIOGRAPHY.

[1] P. Erdös, "On the least primitive root of a prime," *Bulletin of the American Mathematical Society*, vol. 51 (1945), pp. 131-132.

[2] L. R. Ford, *Automorphic Functions*, New York, 1929.

[3] H. Frasch, "Die Erzeugenden der Hauptkongruenzgruppen für Primzahlstufen," *Mathematische Annalen*, vol. 108 (1933), pp. 229-252.

[4] R. Fueter, "Ueber primitive Wurzeln von Primzahlen," *Commentarii Mathematicae Helv.*, vol. 18 (1946), pp. 217-223.

[5] E. Grosswald, "On the genus of the fundamental region of some subgroups of the modular group," *American Journal of Mathematics*, vol. 74 (1952), pp. 86-88.

[6] L. K. Hua, "On the least primitive root of a prime," *Bulletin of the American Mathematical Society*, vol. 48 (1942), pp. 726-730.

[7] J. Nielsen, "A study concerning the congruence subgroups of the modular group," *Det Kgl. Danske Videnskabernes Selskab. Matematisk-Fysiske Meddelelsen*, vol. 25, no. 18 (1950).

[8] S. Pillai, "On the smallest primitive root of a prime," *Journal of the Indian Mathematical Society* (New Series), vol. 8 (1944), pp. 14-17.

[9] H. Rademacher, "Ueber die Erzeugenden von Kongruenzuntergruppen der Modulgruppe," *Abhandlungen aus dem Seminar der Hamburgischen Universität*, vol. 7 (1929), pp. 134-138.

ON SOLUTIONS OF LINEAR PARTIAL DIFFERENTIAL EQUATIONS OF MIXED TYPE.*

By STEFAN BERGMAN.

1. Integral operators in the theory of linear partial differential equations. The realization that a number of relations between some properties of solutions of a homogeneous linear partial differential equation [1]

$$\boldsymbol{L}(\psi) = \psi_{z\bar{z}} + A\psi_z + B\psi_{\bar{z}} + F\psi = 0, \tag{1.1}$$

$z = x + iy, \bar{z} = x - iy,$

$$(\partial\psi/\partial z) = (1/2)[(\partial\psi/\partial x) - (\partial\psi/\partial y)],$$
$$(\partial\psi/\partial\bar{z}) = (1/2)[(\partial\psi/\partial x) + i(\partial\psi/\partial y)],$$

in the case where A, B, F are entire functions of their arguments, are either independent of the coefficients A, B, F of the equation $\boldsymbol{L}$, or depend only upon some properties of A, B, F, has been one of the most significant achievements of the theory of integral operators of the first kind. See [3, 5, 6].[2]

By the operation

$$\psi = \boldsymbol{O}(f) = \mathrm{Re}[\mathrm{P}(f; z_0)]; \tag{1.2}$$

$$\mathrm{P}(f; z_0) = \int_{t=-1}^{+1} E(z', \bar{z}', t) f(z'(1-t^2)) dt/(1-t^2)^{1/2}$$
$$z' = z - z_0, \bar{z}' = \bar{z} - \bar{z}_0, \qquad \mathrm{Re} = \text{Real part},$$

where $E(z', \bar{z}', t)$ (the generating function of the integral operator) is a fixed function, an arbitrary function f of one complex variable is transformed into a solution ψ of (1.1).

Remark. The point (z_0), which appears as an arbitrary parameter in the

* Received March 7, 1951. Research paper done under Navy Contract Nord 10-449, Task 3.

The ideas expressed in this paper represent the personal views of the author, and are not necessarily those of the Bureau of Ordnance.

[1] For our purposes, it is convenient to use complex notation and sometimes to continue x, y to complex values of the arguments, even if our final results refer to the behavior of the solutions in the real domain, i. e., for real values of x and y.

[2] The numbers in brackets refer to the bibliography. A familiarity with previous publications except some portions of [9] is not assumed in the present paper.

operator, will be called its *point of reference.* Usually we set $z_0 = \bar{z}_0 = 0$, unless, as in the case in section 5, it is of essential importance to have $z_0, \bar{z}_0$ arbitrary. The possibility of using an arbitrary point (z_0) as a reference point becomes important if we consider functions f which are singular at the origin, or if the coefficients of the equation become singular there. (This is the case, for instance, for the equations considered subsequently).

To every equation (1. 1), it is possible to find a function $E = E_1$ (the *generating function of the first kind*) which is determined in a unique way, so that the operator $\boldsymbol{O}_1 = \mathrm{Re}[\mathrm{P}_1]$ preserves many properties of the operator "Re."

Using this approach, it has been possible to show that various chapters of the theory of functions of a complex variable can be interpreted not only as theorems on harmonic functions, but also as theorems on solutions of equation (1. 1) with entire coefficients A, B, F.

In particular [3] it has been shown that relations between the properties of the subsequence $\{A_{n0}\}$, $n = 0, 1, 2, \cdots$, of the coefficients of the series development $\psi = \sum_{m,n=0}^{\infty} A_{mn} z^m \bar{z}^n$ of a real solution ψ of $\boldsymbol{L}(\psi) = 0$, and the location and properties of its singularities are independent of the coefficients A, B, F, provided that these latter functions are entire in z and $\bar{z}$.

It is of interest that the method of transforming functions of one variable into solutions of linear differential equations can be extended to certain equations of mixed type. This is of particular importance for the theory of compressible fluids, which is based on the study of equations of this type. In this theory we investigate the equation

$$\boldsymbol{M}(\psi) \equiv \psi_{\mathrm{HH}} + l(\mathrm{H})\psi_{\theta\theta} = 0, \tag{1.3}$$

$$l(\mathrm{H}) = \sum_{n=1}^{\infty} a_n(-\mathrm{H})^n, \quad a_1 > 0, \quad a_2 < 0, \tag{1.4}$$

where $l(\mathrm{H}) > 0$ for $\mathrm{H} < 0$, and $l(\mathrm{H}) < 0$ for $\mathrm{H} > 0$. Reducing (1. 3) to the normal form, one obtains the equation

$$\boldsymbol{L}(\psi) \equiv \psi_{z\bar{z}} + N(\lambda)(\psi_z + \psi_{\bar{z}}) = 0, \quad N = (1/8)l^{-3/2}l_{\mathrm{H}}, \quad \bar{Z} = \lambda - i\theta, \tag{1.1a}$$

$$-\lambda(-\mathrm{H}) = \int_{t=0}^{-\mathrm{H}} [l(t)]^{-1/2} dt, \tag{1.5}$$

[3] For other properties such as the existence of classes analogous to polynomials, approximation developments, summation theorems, the reader is referred to [3, 5, 6].

where $N(\lambda)$ is now singular [4] for $\lambda = 0$. Assuming that l and therefore N is an analytic function of the variable λ, it is convenient to continue our functions to complex values of the arguments. The domains $[-\infty < \theta < \infty, H < 0]$ and $[-\infty < \theta < \infty, H > 0]$ in the real plane correspond to the domains $[-\infty < \theta < \infty, \operatorname{Re}\lambda_1 < 0, \operatorname{Im}\lambda_1 = 0]$ and $[-\infty < \theta < \infty, \operatorname{Im}\lambda_1 > 0, \operatorname{Re}\lambda_1 = 0]$ respectively. Here $\lambda_1 = \lambda + i\Lambda$.

Since the integral operator of the first kind is not defined at the points where the coefficient N becomes singular, it is useful, in studying (1. 1a) to introduce two *integral operators* $\mathbf{P}_{21}$ and $\mathbf{P}_{22}$ (see (5. 8), (4. 1), and (4. 2)) *of the second kind,* which are defined in two different domains of the λ, θ-plane. Each of these operators is defined only in subdomains of the regularity domain of N, but $\mathbf{P}_{22}$ is defined also on the singularity line $\lambda = 0$ of N. $\mathbf{P}_{21}$ can be defined [5] for $\lambda \to -\infty$.

As is the case of regular coefficients A, B, F, it is possible to show that for a class of equations (1. 1a), namely those for which $N(\lambda)$ satisfies the conditions (3. 5a)-(3. 5d), a number of relations are either independent of $N(\lambda)$ or depend only on some properties of $N(\lambda)$. Among equations of this type, it is possible to select one, to which we shall refer as the "simplified case," where the situation is particularly simple, and whose theory can be used as a pattern for the development of the theory in the general case of equations of this type. In the "simplified case" where $l(H) = -a_1 H$, a_1 a positive constant, the generating function [6] $E_2\dagger$ of the integral operator of the second kind is given by

$$(1.6) \qquad E_2\dagger = H\dagger E_2{}^*\dagger(Z, \bar{Z}, t), \quad H\dagger = S_0(-2\lambda)^{-1/6}, \quad Z = \lambda + i\theta,$$

where $E_2{}^*\dagger$ is a hypergeometric function of the variable $u = t^2 Z/2\lambda$,

$$(1.7) \qquad E_2{}^*\dagger = A_1 F(1/6, 5/6, 1/2, u) + B_1 u^{1/2} F(2/3, 4/3, 3/2, u).$$

A_1 and B_1 are constants.

By the transformation $E_2{}^* = H^{-1}E_2$, equation (3. 2) of [9] for E_2 is reduced to the somewhat simpler equation (4. 9) for $E_2{}^*$.

---

[4] The study of the equations of mixed type is in this and many other respects equivalent to the study of partial differential equations of the form (1. 1) with singular coefficients.

[5] These results are of importance especially in fluid dynamics. To $\lambda = -\infty$ corresponds in the hodograph plane (i. e., in the plane whose Cartesian coordinates are the components u, v of the velocity) the origin $u = v = 0$.

[6] The generating function and some other quantities referring to the simplified case will be denoted by symobls with a dagger (†).

It has been shown in the general case that E_2* satisfies a *partial differential equation* (see 4. 9), which in the " simplified case " becomes an *ordinary* differential equation (the hypergeometric equation). Equation (4. 9) possesses some features which can be considered a generalization of properties of hypergeometric equation.

In particular E_2*† can be developed into a series of powers of u (and $v = 1 - u$), converging in $|Z/2\lambda| < 1$, as well as into a series of powers of u^{-1} (and $(1 - v^{-1})$) converging in $|Z/2\lambda| > 1$. Continuing the investigations of [9] in sections **4** and **6**, we show that in the general case of equation (1. 1a) there exist generating functions E_{21} and E_{22} representable in the form of infinite series, converging in $|Z/2\lambda| < 1$ and in $|Z/2\lambda| > 1$, respectively. They can be considered as generalizations of the above-mentioned series developments for the hypergeometric function. Substituting E_{21} and E_{22} into (1. 2) instead of E, we define the integral operators $\mathbf{P}_{21}$ and $\mathbf{P}_{22}$ in $|Z/2\lambda| < 1$ and in $|Z/2\lambda| > 1$, respectively.

Thus in the course of our investigations, we introduce a variety of integral operators which are defined in different domains, and which preserve different properties of the function f to which they are applied. In order to be able to employ different operators simultaneously, however, we must determine relations which exist between the associate functions of the same solution $\psi(\lambda, \theta)$ of (1. 1a), but with respect to different integral operators. In section **5** we determine the associate of $\mathbf{P}_{21}$. This result solves the *initial value problem in the large*, with the data $\psi(\lambda, \theta)$, $\psi_\lambda(\lambda, \theta)$ given on a segment $\lambda = \lambda_0$, $\lambda_0 < 0$, i. e., in the domain of *elliptic* behavior. In section **7** a similar problem is considered with $\psi(0, \theta)$, $\lim_{\lambda \to 0}(-\lambda)^{1/3}\psi_\lambda(\lambda, \theta) = C\psi_H(H, \theta)_{H=0}$ (see footnote [22] on p. 468) given, i. e. with data on the transition line, $\lambda = 0$. Our results give information about the behavior of the solution in regions lying in the domain of validity of the corresponding operator. We also clarify the extent to which these results depend upon the coefficient N of equation (1. 1a).

Solutions of linear differential equations with analytic coefficients are analytic functions of two variables. When the ψ's are continued to complex values of the arguments, they represent important subclasses of functions of two complex variables. As indicated in section **8**, the results obtained in sections **5** and **7** can be interpreted from this point of view. They complement in a natural way that portion of the theory of functions of two (and several) complex variables which is devoted to the study of the relations between the coefficients of the series development of a function, its growth and value distribution.

The theorems of sections **5** and **7** illustrate the application of the operator method. The results of the present paper can be employed for the study of various other questions in the theory of differential equations of the type (1. 3), in a manner similar to that of the case of regular coefficients.

The author wishes to express his thanks to Dr. Philip Davis for his assistance in the preparation of this paper and various valuable suggestions which have been incorporated.

2. Solution of the initial value problem in the small. In this section we shall, by using a set of particular solutions introduced by the author[7] (cf. [4]), solve the initial value problem *in the small* for equation (1. 3) with the data given on the transition line.

The above-mentioned solutions are

$$(2.1)\quad \psi_{n1}(\mathrm{H},\theta) = 1!\binom{n}{1}\theta^{n-1}\int_0^{\mathrm{H}} d\mathrm{H} - 3!\binom{n}{3}\theta^{n-3}\int_0^{\mathrm{H}} d\mathrm{H}_3\int_0^{\mathrm{H}} l(\mathrm{H}_2)d\mathrm{H}_2\int_0^{\mathrm{H}_2} d\mathrm{H} + \cdots,$$

$$(2.2)\quad \psi_{n2}(\mathrm{H},\theta) = \theta^n - 2!\binom{n}{2}\theta^{n-2}\int_0^{\mathrm{H}} d\mathrm{H}_2\int_0^{\mathrm{H}_2} l(\mathrm{H}_1)d\mathrm{H} + \cdots.$$

In this section, we assume that $l(\mathrm{H})$ is bounded in some interval $-\mathrm{H}' \leq \mathrm{H} \leq \mathrm{H}'$. From this assumption it follows that all the functions ψ_{nk} $(n = 1, 2, \cdots; k = 1, 2)$, are twice-differentiable over $[-\mathrm{H}' \leq \mathrm{H} \leq \mathrm{H}'; -\infty < \theta < \infty]$. These solutions have the further property that

$$(2.3)\quad \psi_{n1}(0,\theta) = 0, \qquad \psi_{n2}(0,\theta) = \theta^n,$$

$$(2.4)\quad [\partial\psi_{n1}(\mathrm{H},\theta)/\partial\mathrm{H}]_{\mathrm{H}=0} = n\theta^{n-1}; \quad [\partial\psi_{n2}(\mathrm{H},\theta)/\partial\mathrm{H}]_{\mathrm{H}=0} = 0.$$

THEOREM 2. 1. *Let the functions* $\chi_k(\theta)$ $(k = 1, 2)$, *possess expansions of the form*

$$(2.6)\quad \chi_k(\theta) = \sum_{\nu=0}^{\infty} a_\nu^{(k)}\theta^\nu, \qquad |\theta| \leq \theta_1.$$

Then

$$(2.7)\quad \psi(\mathrm{H},\theta) = \sum_{n=1}^{\infty} n^{-1}a_{n-1}^{(2)}\psi_{n1}(\mathrm{H},\theta) + \sum_{n=0}^{\infty} a_n^{(1)}\psi_{n2}(\mathrm{H},\theta)$$

---

[7] These functions were introduced independently in a joint paper [12] (see also [13]), by L. Bers and A. Gelbart, who term them σ-monogenic functions.

converges in a domain

$$D = \mathrm{E}[|\theta| + \sigma|\mathrm{H}| < \theta_1]; \quad \sigma = \sigma(l(\mathrm{H})) > 0. \tag{2.8}$$

The series (2.7) *represents a solution of equation* (1.2) *for which*

$$\psi(0, \theta) = \chi_1(\theta), \tag{2.9a}$$

$$[\psi_{\mathrm{H}}(\mathrm{H}, \theta)]_{\mathrm{H}=0} = \chi_2(\theta), \qquad \psi_{\mathrm{H}} = \partial\psi/\partial\mathrm{H}, \tag{2.9b}$$

holds.

Proof. From the convergence of the series in (2.6) it follows that

$$|a_n^{(1)}| \leq \mu_1/\theta_1^n; \qquad |a_2^{(n)}| \leq \mu_2/\theta_1^n, \qquad (n = 0, 1, 2, \cdots), \tag{2.10}$$

for some conveniently chosen positive constants μ_1 and μ_2. Inasmuch as $l(\mathrm{H})$ is assumed to be bounded, we have for some positive constant c,

$$|l(\mathrm{H})| \leq c^2 \qquad \text{for } |\mathrm{H}| \leq \mathrm{H}'. \tag{2.11}$$

Now select

$$\sigma > \max[\theta_1/\mathrm{H}', c]. \tag{2.12}$$

From (2.1), (2.2), and (2.11), it follows that

$$|\psi_{nk}(\mathrm{H}, \theta)| \leq (|\theta| + c|\mathrm{H}|)^n, \qquad \text{for } |\mathrm{H}| \leq \mathrm{H}'. \tag{2.13}$$

If $(\mathrm{H}, \theta) \in D$, then $|\mathrm{H}| \leq (\theta_1/\sigma) < \mathrm{H}'$ by (2.12), so that (2.13) holds for (H, θ) in this domain *a fortiori*. Thus, for $(\mathrm{H}, \theta) \in D$, the series (2.7) is dominated by

$$(\mu_1 + \mu_2) \sum_{n=0}^{\infty} [(|\theta| + c|\mathrm{H}|)/\theta_1]^n. \tag{2.14}$$

But by (2.12) we have $|\theta| + c|\mathrm{H}| < \theta_1$ for $(\mathrm{H}, \theta) \in D$, and therefore (2.7) converges absolutely and uniformly in D. By a similar argument, it is clear that the first two derived series also converge absolutely and uniformly in D, and therefore (2.7) is a solution of (1.1a). From (2.3) and (2.4) it follows that this solution satisfies the initial conditions (2.9).

Remark. Let A be a class of functions of H and θ which are infinitely differentiable, and which possess the further property that any function of A is uniquely determined by the derivatives $[\partial^{m+n}\psi(\mathrm{H}, \theta)/\partial\mathrm{H}^m\partial\theta^n]_{\mathrm{H}=0, \theta=0}$ $(m, n = 0, 1, 2, \cdots)$ (as, for example, the class of functions analytic in a neighborhood of $\mathrm{H} = 0$, $\theta = 0$). If $l(\mathrm{H})$ is infinitely differentiable over $-\mathrm{H}' \leq \mathrm{H} \leq \mathrm{H}'$, and the solution (2.9) of the partial differential equation

is required to belong to a class A, then it is unique. For since $\psi(0,\theta)=\chi_1(\theta)$, and $\psi_{\mathrm{H}}(0,\theta)=\chi_2(\theta)$, it follows that $(\partial^m\psi(0,\theta)/\partial\theta^m)_{\theta=0}=\chi_1^{(m)}(0)$, and $[\partial^{m+1}\psi(\mathrm{H},\theta)/\partial\mathrm{H}\partial\theta^m]_{\mathrm{H}=0,\theta=0}=\chi_2^{(m)}(0)$ $(m=0,1,2,\cdots)$, while the remaining mixed derivatives may be obtained from $\chi_i(\theta)$ and $l(\mathrm{H})$ by recursion through the repeated differentiation of $\psi_{\mathrm{HH}}+l(\mathrm{H})\psi_{\theta\theta}=0$.

3. Reduction of the equation of mixed type to normal form; the class $\mathfrak{N}$ of differential equations of type (3.1). The function $l(\mathrm{H})$ given by (1.3) will now be assumed to satisfy some additional conditions, the precise formulation of which will be given below. Our aim is to show that certain properties of solutions of (1.1) are independent of the special form of $l(\mathrm{H})$, provided that $l(\mathrm{H})$ satisfies these conditions.

In order to formulate these conditions, it is convenient to reduce (1.1) to the normal form

$$\psi_{\lambda\lambda}+\psi_{\theta\theta}+4N(\lambda)\psi_\lambda=0; \qquad N=(1/8)l^{-3/2}l_{\mathrm{H}}, \tag{3.1}$$

by means of the transformation

$$-\lambda(-\mathrm{H})=\int_{t=0}^{-\mathrm{H}}[l(t)]^{-1/2}dt. \tag{3.2}$$

A formal computation shows that $-\lambda(-\mathrm{H})$ has an expansion of the form

$$-\lambda(-\mathrm{H})=(2/3)a_1^{1/2}(-\mathrm{H})^{3/2}+(1/3)a_2a_1^{-1/2}(-\mathrm{H})^{5/2}+\cdots, \tag{3.3}$$

valid in a neighborhood of $\mathrm{H}=0$, while N has an expansion of the form

$$N(\lambda)=(-\lambda)^{-1}[-(1/12)+\sum_{\nu=1}^{\infty}\beta_\nu(-\lambda)^{2\nu/3}], \qquad \beta_1>0, \tag{3.4}$$

valid for $-\lambda_0<\lambda<0$, $\lambda_0>0$. If r_0 is the radius of convergence of the series $\sum_{\nu=1}^{\infty}\beta_\nu x^\nu$, then we have $\lambda_0=r_0^{3/2}$.

We now impose the following requirements[8] on the function $N(\lambda)$. We shall say that a function $N(\lambda)$ of the real variable λ belongs to the class $\mathfrak{N}$, if

(3.5a) $N(\lambda)$ is an analytic function for $-\infty<\lambda<0$, and is real for $\lambda<0$;

[8] These requirements, of course, may be replaced so as to give conditions directly on the function $l(\mathrm{H})$; however, such a formulation would be more involved than that given above.

(3.5b) in the neighborhood of $\lambda = 0$, $N(\lambda)$ possesses an expansion of the form (3.4), valid for $-\lambda_0 < \lambda < 0$; $0 < \lambda_0 \leqq \infty$;

(3.5c) the integral [9] $\exp[-\int_{-\infty}^{\lambda} 2N(t)\,dt] = H(\lambda)$, exists for all $\lambda < 0$.

Under these assumptions, it may be shown that $\lim_{\lambda \to 0^-} (-\lambda)^{1/6} H(\lambda)$ exists.[10] Set S_0 equal to the value of this limit. We shall assume further [11] that

$$S_0 > 0. \tag{3.5d}$$

Remark. If (3.1) possesses a coefficient $N(\lambda)\,\varepsilon\,\mathfrak{N}$, we shall say that the differential equation is of class $\mathfrak{N}$.

Before investigating the general equation (3.1), it is useful to consider the special case

$$l(H) = -a_1 H, \qquad a_1 > 0, \tag{3.6}$$

to which, following [9], p. 862, we shall refer as "the simplified case." In this case $N(\lambda) = (12\lambda)^{-1}$. If the lower limit appearing in the integral in (3.5c) is replaced by any a for which $-\infty < a < 0$, it will be found that conditions (3.5a)-(3.5c) are satisfied. In the simplified case, the generating

---

[9] In compressible fluids this integral equals $l^{-1/4} = (\rho^2(1-M^2))^{1/4}$.

[10] The existence of S_0 may be established as follows. Select λ_1 such that $-\lambda_0 < \lambda_1 < 0$. Then $H(\lambda) = \exp[\int_{-\infty}^{\lambda_1} (-2N(t))\,dt + \int_{\lambda_1}^{\lambda} (-2N(t))\,dt]$. Introduce

$$T(\lambda) = -(1/6)\log(-\lambda) + \sum_{\nu=1}^{\infty} (6\beta_\nu/(2\nu-3))(-\lambda)^{2\nu/3},$$

$(-\lambda_0 < \lambda < 0)$; then (independently of λ_1),

$$H(\lambda) = \exp[\int_{-\infty}^{\lambda_1} [-2N(t)]\,dt + T(\lambda) - T(\lambda_1)]$$
$$= \exp[\int_{-\infty}^{\lambda_1} (-2N(t))\,dt - T(\lambda_1)](-\lambda)^{-(1/6)}[1 + \cdots].$$

Thus

$$S_0 = \exp[\int_{-\infty}^{\lambda_1} (-2N(t))\,dt - T(\lambda_1)].$$

It should be pointed out that (4.3) of [9] should be corrected to read as (3.5c) above. Also on p. 860 of [9], line 4 from the bottom, "i" should be omitted. On p. 871, lines 12 and 11, $E_z{}^*\dagger$, $E_{zt}{}^*\dagger$, and $E_{zz}{}^*\dagger$ should be replaced by $E_{\bar{z}}{}^*\dagger$, $E_{\bar{z}t}{}^*\dagger$, and $E_{\bar{z}\bar{z}}{}^*\dagger$. On p. 872, formula (4.9'), the exponent "$-2/3$" should be replaced by "$2/3$."

[11] The value of S_0 in the case of compressible fluids is given in (4.3) of [9]. A detailed derivation of this formula can be found in the appendix p. 898 to the Technical Report No. 10 of the series "Operator Methods in the Theory of Compressible Fluids," Harvard University.

function of the second kind $E\dagger$ of the integral operator is the hypergeometric function of the variable $u = Zt^2/(Z+\bar{Z}) = t^2Z/2\lambda$, $Z = \lambda + i\theta$, $\bar{Z} = \lambda - i\theta$;

$$(3.7) \qquad E_2{}^*\dagger = A_1F(1/6, 5/6, 1/2, u) + B_1u^{1/2}F(2/3, 4/3, 3/2, u),$$

and it is possible to represent it in the form

$$(3.8) \qquad E_{21}{}^*\dagger^{(k)} = \sum_{n=0}^{\infty} (t^2Z)^{n-(1/2)+(k/2)}Q\dagger^{(n,k)}(2\lambda) \qquad (k=1,2),$$

holding for $|Z/2\lambda| < 1$, and

$$(3.9) \qquad E_{22}{}^*\dagger^{(k)} = \sum_{n=0}^{\infty} q\dagger^{(nk)}(2\lambda)/(-t^2Z)^{n-(1/2)+(2k/3)} \qquad (k=1,2),$$

holding for $|Z/2\lambda| > 1$, where [12]

$$(3.10) \quad Q\dagger^{(n1)}(2\lambda) = \frac{(1/6)[(1/6)+1]\cdots[(1/6)+n-1](5/6)[(5/6)+1]\cdots[(5/6)+n-1]}{n!(1/2)[(1/2)+1]\cdots[(1/2)+n-1](2\lambda)^n}$$

$$Q\dagger^{(01)}(2\lambda) = 1,$$

and where [13]

$$(3.11) \quad q\dagger^{(n1)}(\lambda) = \frac{(1/6)[(1/6)+1]\cdots[(1/6)+n-1](2/3)[(2/3)+1]\cdots[(2/3)+n-1](-2\lambda)^n}{n!(1/3)[(1/3)+1]\cdots[(1/3)+n-1]}$$

and

$$(3.12) \quad q\dagger^{(n2)}(\lambda) = \frac{(5/6)[(5/6)+1]\cdots[(5/6)+n-1](4/3)[(4/3)+1]\cdots[(4/3)+n-1](-2\lambda)^{n+1}}{n!(5/3)[(5/3)+1]\cdots[(5/3)+n-1]}$$

4. The integral operator of the second kind for the general equation. In this section we shall show that for the general case there exist two series developments which are analogous to (3.8) and (3.9). They may be written in the form [14]

---

[12] $Q^{(n,2)}(2\lambda)$ has not been listed here inasmuch as the corresponding operator, after integration, yields zero identically.

[13] Formula (5.22) of [9] should be corrected as above. Formula (4.14) of [9] should read "$+(b_3S_0/(2\Lambda)^{1/6})(2\Lambda/(2\Lambda - t^2(\Lambda+\theta)))^{1/2}F(\cdots$" On p. 880, formula (5.20) should read "for $2|\lambda| < Z$, and $|\lambda|^{2/3} < \varepsilon_0$."

[14] In addition to developments in powers of u and $(1/u)$, the hypergeometric series has other representations. In section 5 we shall derive representations corresponding to developments in powers of $1-u$ and $1-(1/u)$.

$$E_{21}{}^{*(k)} = \sum_{n=0}^{\infty} (t^2Z)^{n-(1/2)+k/2} Q^{(nk)}(2\lambda) \qquad (k=1,2), \tag{4.1}$$

$$(\lambda < 0, |\theta|^2 < 3|\lambda|^2),$$

and

$$E_{22}{}^{*(k)} = \sum_{n=0}^{\infty} q^{(nk)}(2\lambda)/(-t^2Z)^{n-(1/2)+2k/3} \qquad (k=1,2), \tag{4.2}$$

$$(|\theta|^2 > 3|\lambda|^2, |\lambda| < \lambda_0),$$

where [15]

$$(2n+1)Q_\lambda{}^{(n+1\,1)} + Q_{\lambda\lambda}{}^{(n1)} + 4FQ^{(n1)} = 0, \qquad Q^{(n+1\,1)}(-\infty) = 0, \tag{4.3}$$

$$Q_\lambda{}^{(11)} = -4F; \qquad F(\lambda) = -N^2(\lambda) - (1/2)N_\lambda(\lambda)$$

$$Q^{(0k)} = (2\lambda)^{(1/2)-k/2},$$

while the $q^{(nk)}$ are given by the recursion relations

$$q_{\lambda\lambda}{}^{(0k)} + 4F(\lambda)^{(0k)} = 0 \qquad (k=1,2), \tag{4.4}$$

$$2[n + (2k/3)]q_\lambda{}^{(nk)} + q_{\lambda\lambda}{}^{(n+1\,k)} + 4F(\lambda)q^{(n+1\,k)} = 0,$$

and satisfy the initial conditions which are implicit in (4.5).

THEOREM 4.1. *Let $q^{(nk)}(\lambda)$ $(n=0,1,2,\cdots; k=1,2)$, be a set of functions which satisfy the relations* (4.4) *and in the neighborhood of the origin have developments of the form*

$$q^{(nk)}(\lambda) = \sum_{\nu=0}^{\infty} C_\nu{}^{(nk)}(-\lambda)^{n-(1/2)+(2/3)(k+\nu)}, \tag{4.5}$$

where

$$C_0{}^{(01)} = 2^{1/6}, \qquad C_0{}^{(n1)} = q\dagger^{(n1)}(-1), \tag{4.6a}$$

$$C_1{}^{(n1)} = 0, \tag{4.6b}$$

$$C_0{}^{(02)} = 2^{5/6}, \qquad C_0{}^{(n2)} = q\dagger^{(n2)}(-1). \tag{4.6c}$$

Then the series

$$E_{22}{}^{*(k)} = \sum_{n=0}^{\infty} \frac{q^{(nk)}(\lambda)}{(-t^2Z)^{n-(1/2)+(2k/3)}} \qquad (k=1,2), \tag{4.7}$$

converges uniformly and absolutely for Z belonging to any closed subdomain of

$$\mathrm{E}[2|\lambda| < |Z|, \lambda \leq 0] \quad \text{and} \quad |t| \geq 1. \tag{4.8}$$

[15] The conditions $Q^{(n+1\,1)}(a) = 0$, $a > -\infty$, could also have been imposed.

The series (4.7) *then represents a solution of the partial differential equation*

$$E_{Zt}^* - t^2 E_{\bar{Z}t}^* + 2Zt\,FE_{Z\bar{Z}}^* + 2Zt\,FE^* = 0, \tag{4.9}$$

where

$$Z = \lambda + i\theta, \qquad \bar{Z} = \lambda - i\theta, \qquad \lambda = (Z + \bar{Z})/2, \tag{4.10}$$

and is, therefore, when multiplied by $H(\lambda)$, *a generating function of the second kind defined in the domain* (4.8) *for the differential equation* (3.1).

Proof. As has been shown[16] in [9], the series (4.7) converges absolutely and uniformly for $|t| \geq 1$ and for λ, Z belonging to the region $\mathbf{E}[2|\lambda| < |Z|, \lambda < 0, |-\lambda| < \lambda_0]$, and represents a solution of the differential equation (4.9). It remains to show that the convergence holds for $\lambda < -\lambda_0$. We shall first establish several lemmas. It is convenient to introduce the variable $s = (-\lambda)^{2/3}$, noting that $-\lambda$ is positive and choosing the real branch. In this variable, we shall have

$$F(\lambda(s)) = s^{-3}S(s), \tag{4.11}$$

where $S(s)$ is regular for $|s| < |\lambda_0|^{2/3}$. Moreover, in this variable, the relations (4.4) become

$$s^2 q_{ss}^{(0k)} - (1/2)sq_s^{(0k)} + 9S(s)q^{(0k)} = 0, \tag{4.12a}$$

$$-3[n + (2k/3)]s^{5/2}q_s^{(nk)} + s^2 q_{ss}^{(n+1\,k)} - (1/2)sq_s^{(n+1\,k)} + 9S(s)q^{(n+1\,k)} = 0. \tag{4.12b}$$

REMARK. We note that for $q(\lambda)$ and $q(\lambda(s))$ the same symbol q is used.

LEMMA 1. *Let*

$$w^{(k)}(s) = s^{[k-(3/4)]} \sum_{\nu=0}^{\infty} C_\nu^{(0k)} s^\nu \qquad (k = 1, 2). \tag{4.13}$$

Then we have

$$q^{(n+1\,k)}(s) = 3[n + (2k/3)]w^{(k)}(s) \int_0^s [w^{(k)}]^{-2} s_1^{(1/2)} \left(\int_0^{s_1} w^{(k)} q_{s_2}^{(nk)} ds_2 \right) ds_1,$$
$$(n = 0, 1, 2, \cdots; k = 1, 2).$$

[16] The special case of this theorem where N is that which appears in the theory of compressible fluids, i.e., when it is given by (2.16), has already been stated on p. 878 of [9]. However, the proof given there is only for λ belonging to the domain $\mathbf{E}[2|\lambda| < |Z|, -\lambda_0 < \lambda \leq 0]$. The generalization of this part of the proof to the general case, i.e. $N(\lambda) \in \mathfrak{R}$, is immediate; the main content of the present section is the removal of the restriction $-\lambda_0 < \lambda \leq 0$. For λ_0, see (3.4).

Proof. It has been shown on page 879 of [9] that $w^{(k)}(s) = q^{(0k)}(s)$, so that by (4.12a), $w^{(k)}(s)$ $(k = 1, 2)$ are two independent solutions of the differential equation

$$s^2 w_{ss} - (1/2) s w_s + 9S(s) w = 0. \tag{4.14}$$

Introduce

$$u^{(n+1k)}(s) = 3[n + (2k/3)] \int_0^s [w^{(k)}]^{-2} s_1^{(1/2)} \left(\int_0^{s_1} w^{(k)} q_{s_2}^{(nk)} ds_2 \right) ds_1, \tag{4.15}$$

so that it is necessary to prove that

$$q^{(n+1k)}(s) = w^{(k)}(s) u^{(n+1k)}(s), \qquad n = -1, 0, 1, 2, \cdots. \tag{4.16}$$

Substituting the right-hand side of (4.16) into (4.12b), we obtain

$$\begin{aligned} s^2 w^{(k)} u_{ss}^{(n+1k)} + u_s^{(n+1k)} [2s^2 w_s^{(k)} - (1/2) s w^{(k)}] \\ = 3[n + (2k/3)] s^{5/2} q_s^{(nk)}, \end{aligned} \tag{4.17}$$

or

$$\begin{aligned} w^{(k)} u_{ss}^{(n+1k)} + u_s^{(n+1k)} [2 w_s^{(k)} - 2^{-1} s^{-1} w^{(k)}] \\ = 3[n + (2k/3)] s^{1/2} q_s^{(nk)}. \end{aligned} \tag{4.18}$$

Now it may be easily verified by differentiation that (4.15) is a particular solution of (4.18), and therefore that (4.16) is a solution of (4.12b).

The first terms in the expansion of $w^{(k)}(s) u^{(n+1k)}(s)$ are given by

$$\begin{aligned} &\frac{[(3n/2) - (3/4) + k](3n + 2k) C_0^{(nk)}}{[(3n/2) + 2k - (3/2)][(3n/2) + (3/2)]} s^{[(3n/2)+k+(3/4)]} \\ &\quad + \left\{ \frac{[(3n/2) + (1/4) + k](3n + 2k) C_1^{(nk)}}{[(3n/2) + 2k - (1/2)][(3n/2) + (5/2)]} \right. \\ &\quad + \frac{(3n + 2k) C_1^{(0k)} C_0^{(nk)}}{C_0^{(0k)}} \left[\frac{[(3n/2) - (3/4) + k]}{[(3n/2) - (1/2) + 2k][(3n/2) + (5/2)]} \right. \\ &\quad - \frac{2[(3n/2) - (3/4) + k]}{[(3n/2) - (3/2) + 2k][(3n/2) + (5/2)]} \\ &\left. \left. + \frac{[(3n/2) - (3/4) + k]}{[(3n/2) + (3/2)][(3n/2) + 2k - (3/2)][(3n/2) + (5/2)]} \right] \right\} s^{[(3n/2)+k+(7/4)]} + \cdots. \end{aligned}$$

(See (4.13), (4.5) and (4.15)). From the coefficients in this expansion, it may be verified by induction that the conditions (4.6) are satisfied. Thus we have shown that $q^{(n+1k)}(s)$ and $w^{(k)}(s) u^{(n+1k)}(s)$ are both solutions of (4.12b) and are such that their first two coefficients coincide. Our lemma now follows.

LEMMA 2. *The functions*

$$w_0 = \exp\left[\int_{-\infty}^{\lambda} 2N d\lambda\right] = (1/H), \tag{4.19}$$

(see (3.5c)) *and*

$$w_2 = -2^{11/6}(3S_0H)^{-1}\int_0^{\lambda} H^2 d\lambda, \tag{4.20}$$

$$S_0 = \lim_{\lambda\to 0^-} (-\lambda)^{1/6}\exp\left[-2\int_{-\infty}^{\lambda} N(t)\,dt\right],$$

both satisfy the differential equation (4.14).

Proof. We have

$$\lambda = -s^{3/2}, \qquad d\lambda/ds = -(3/2)s^{1/2}, \qquad d^2\lambda/ds^2 = -(3/4)s^{-1/2}. \tag{4.21}$$

so that

$$dw_0/ds = -3s^{1/2}w_0N, \tag{4.22}$$

and

$$d^2w_0/ds^2 = 9sw_0N^2 + (9/2)sw_0N_\lambda - (3/2)s^{-1/2}w_0N. \tag{4.23}$$

Substituting these values in (4.14) and simplifying them, we obtain

$$2N_\lambda + 4N^2 + 4F = 0, \tag{4.24}$$

which, by condition (4.3), is satisfied identically by N.

To prove that (4.20) is a solution of (4.14), we may proceed as follows. Let

$$w_2 = H^{-1}r. \tag{4.25}$$

Then we obtain the following condition on r in order that (4.14) be satisfied. We have

$$dw_2/ds = (1/H)r_s + r(1/H)_s, \tag{4.26}$$

$$d^2w_2/ds^2 = (1/H)r_{ss} + 2r_s(1/H)_s + r(1/H)_{ss},$$

so that (4.14) becomes

$$s^2[(r_{ss}/H) - 2r_s(H_s/H^2) + r(1/H)_{ss}] \tag{4.27}$$

$$-(s/2)[(r_s/H) + r(1/H)_s] + 9[S(s)](r/H) = 0,$$

or

$$s(r_{ss} - 2r_s(H_s/H)) - (1/2)r_s = 0, \tag{4.28}$$

since $(1/H)$ is a solution of (4. 14). The quantity r_s must therefore satisfy

(4. 29) $(r_{ss}/r_s) = (2H_s/H) + (1/2s)$, or

$$(d\log(r_s)/ds) = (2d\log H/ds) + 1/2(d\log s/ds),$$

so that $r_s = cH^2s^{1/2}$, and $r = c\int_0^s s^{1/2}H^2ds = -(2/3)c\int_0^\lambda H^2d\lambda$.

LEMMA 3.

$$w^{(2)}(s) = w_2(s), \tag{4. 30}$$

$$w^{(1)}(s) = 2^{1/6}S_0w_0(s) + 2^{-2/3}S_1w_2(s), \tag{4. 31}$$

where $S_1 = 3\beta_1$, and S_0 has already been given in (3. 5c); and hence

$$w^{(1)}(s) = 2^{1/6}S_0H^{-1} + 2^{1/6}S_1S_0^{-1}H^{-1}\int_0^s H^2s^{1/2}ds, \tag{4. 32}$$

$$w^{(2)}(s) = 2^{5/6}S_0^{-1}H^{-1}\int_0^s H^2s^{1/2}ds. \tag{4. 33}$$

Proof. As we have shown, both sides of (4. 30) and (4. 31) satisfy the linear differential equation (4. 14). Hence, by classical theorems, it suffices to show that the first two coefficients in the expansions coincide. It has been shown in [9], page 870, that in the neighborhood of $\lambda = 0$, $H(\lambda)$ has an expansion of the form

$$H(\lambda) = S_0(-\lambda)^{-1/6} + S_0S_1(-\lambda)^{3/6} + S_0S_2(-\lambda)^{7/6} + \cdots. \tag{4. 34}$$

Hence

$$H^2 = S_0^2(-\lambda)^{-1/3} + \cdots, \quad \int_0^\lambda H^2d\lambda = -(3/2)S_0^2(-\lambda)^{2/3} + \cdots,$$

$$-2^{11/6}(3S_0H)^{-1}\int_0^\lambda H^2d\lambda = 2^{5/6}(-\lambda)^{5/6} + \cdots. \tag{4. 34a}$$

Therefore, from (4. 13), (4. 6c) and (4. 20), since the series development of $w^{(2)}$ and w_2 begins with $s^{5/4}$, both $w^{(2)}$ and w_2 coincide in the first two coefficients $a_1 = 0$ and $a_2 = 2^{5/6}$ of their series development $a_1(-\lambda)^{1/6} + a_2(-\lambda)^{5/6} + \cdots$. We have, furthermore, from (4. 19) and (4. 34),

$$S_0w_0 = (-\lambda)^{1/6} - S_1(-\lambda)^{5/6} + \cdots,$$

and from (4. 20) and (4. 34a) $S_1w_2 = 2^{5/6}S_1(-\lambda)^{5/6} + \cdots$, so that $2^{1/6}S_0w_0 + 2^{-2/3}S_1w_2 = 2^{1/6}(-\lambda)^{1/6} + 0\cdot(-\lambda)^{5/6} + \text{const}(-\lambda)^{9/6} + \cdots$. Therefore, by (4. 13), (4. 6a) and (4. 6b), the first two coefficients of this sum and $w^{(1)}$ coincide.

LEMMA 4. *For an arbitrary $s_2 > 0$ there exist three positive constants $C_1(s_2)$, $C_2(s_2)$, $C_3(s_2)$, such that*

$$C_1 s^{k-(3/4)} \leq w^{(k)}(s) \leq C_2 s^{k-(3/4)}, \text{ for } 0 \leq s \leq s_2 < \infty, \quad (k = 1, 2), \tag{4.35}$$

$$| w_s^{(k)} | \leq C_3 s^{k-(7/4)}, \qquad 0 \leq s \leq s_2. \tag{4.36}$$

Proof. We note that since S_0 and S_1 are both positive, it follows from (4. 19), (4. 32), and (4. 33) that $w^{(k)}(s) \geq 0$, and $w^{(k)}(s) = 0$ only if $s = 0$. By (4. 13),

$$w^{(k)} s^{-k+(3/4)} = \sum_{\nu=0}^{\infty} C_\nu^{(0k)} s^\nu, \qquad | s | < s_0, \tag{4.37}$$

so that $w^{(k)} s^{-k+(3/4)}$ is regular in $| s | < s_0$. Since, in addition, the integrals (4. 32) and (4. 33) are valid for all real $s \geq 0$, the existence of constants for which (4. 35) holds, now follows. Again, since

$$w_s^{(k)} s^{-k+(7/4)} = \sum_{\nu=0}^{\infty} (\nu + k - (3/4)) C_\nu^{(0k)} s^\nu, \qquad | s | < s_0, \tag{4.38}$$

$w_s^{(k)} s^{-k+(7/4)}$ is regular in $| s | < s_0$, so that the existence of C_3 with (4. 36) follows.

We are now in a position to prove our theorem. Recalling that in (4. 15) $u^{(n+1\,k)}(s)$ was defined by

$$u^{(n+1\,k)}(s) = 3[n + (2k/3)] \int_0^s [w^{(k)}]^{-2} s_1^{1/2} \Big(\int_0^{s_1} w^{(k)} q_{s_2}^{(nk)} ds_2 \Big) ds_1, \tag{4.15}$$

we wish to show that for all s_2, $s_2 > 0$, and for all $\epsilon > 0$, there exists a $C = C(s_2, \epsilon)$ such that

$$| u^{(nk)}(s) | \leq C \frac{2^n \Gamma[n + (2k/3)]}{\Gamma(2k/3)\Gamma(n+1)} (1 + \epsilon)^n s^{3n/2}, \qquad 0 \leq s \leq s_2. \tag{4.39}$$

Since, according to page 879 of [9] and (4. 13), $w^{(k)}(s) = q^{(0k)}(s)$, we have, by (4. 16), $u^{(0k)}(s) = 1$, so that statement (4. 39) is obviously true for $n = 0$. From (4. 5), (4. 13), (4. 16), the series development for $u^{(nk)}$ has the form

$$u^{(\mu k)}(s) = C_0^{(\mu k)} [C_0^{(0k)}]^{-1} s^{3\mu/2} \Big(1 + \sum_{\nu=1}^{\infty} a_\nu^{(\mu k)} s^\nu\Big), \qquad 0 \leq s \leq s_0. \tag{4.40}$$

Let ϵ and s_2 be fixed. In view of (4. 40) and the fact that $w^{(k)} \geq 0$ and equals 0 only when $s = 0$, it follows that for every finite n we may determine C so large that the inequality is valid for $\mu = 0, 1, 2, \cdots, n$. For a fixed s_2, choose n so large that

$$C_2 C_3 C_1^{-2} [(3n/2) + 2k - (3/2)]^{-1} < \epsilon, \tag{4.41}$$

where C_1, C_2, C_3 are constants introduced in Lemma 4. Assuming the inequality (4. 39) for $\mu \leq n$, we shall show that it holds for $\mu = n+1$. If we integrate the inner integral in (4. 15) by parts, and note that $w^{(k)}(0) = 0$, we have, by (4. 16),

$$(4.42)\quad u^{(n+1\,k)} = 3[n+(2k/3)] \int_0^s u^{(nk)} s^{(1/2)} ds - 3[n+(2k/3)] \int_0^s (w^k)^{-2} s_1^{1/2} \Big(\int_0^{s_1} w^{(k)} w_{s_2}{}^{(k)} u^{(nk)} ds_2 \Big) ds_1.$$

Therefore, using (4. 35), (4. 36), and (4. 39),

$$\begin{aligned} | u^{(n+1\,k)} | &\leq C \frac{3[n+(2k/3)]2^n \Gamma[n+(2k/3)](1+\epsilon)^n}{\Gamma(2k/3)\Gamma(n+1)} \\ &\quad \times \Big[\frac{s^{3(n+1)/2}}{(3/2)(n+1)} + \frac{C_2 C_3 s^{3(n+1)/2}}{(3/2)C_1{}^2[(3n/2)+2k-(3/2)](n+1)} \Big] \\ &\leq C \frac{2^{n+1}\Gamma[n+1+(2k/3)](1+\epsilon)^n s^{3(n+1)/2}}{\Gamma(2k/3)\Gamma(n+2)} \\ &\quad \times \Big[1 + \frac{C_2 C_3}{C_1{}^2[(3n/2)+2k-(3/2)]} \Big], \end{aligned}$$

and in view of (4. 41),

$$| u^{(n+1\,k)} | \leq \frac{2^{n+1}\Gamma[n+1+(2k/3)](1+\epsilon)^{n+1} s^{3(n+1)/2}}{\Gamma(2k/3)\Gamma(n+2)},$$

which is the inequality for $\mu = n+1$.

Finally, we have

$$(4.43)\qquad E_{22}{}^{*(k)} = \sum_{n=0}^{\infty} \frac{w^{(k)} u^{(nk)}}{(-t^2 Z)^{[n-(1/2)+(2k/3)]}},$$

and

$$\Big| \sum_{n=N}^{\infty} \frac{w^{(k)} u^{(nk)}}{(-t^2 Z)^{[n-(1/2)+(2k/3)]}} \Big| \leq \Big| \frac{w^{(k)}}{(-t^2 Z)^{[(2k/3)-(1/2)]}} \Big| \sum_{n=N}^{\infty} \Big| \frac{u^{(nk)}}{(-t^2 Z)^n} \Big|;$$

by (4. 39),

$$\sum_{n=N}^{\infty} \Big| \frac{u^{(nk)}}{(-t^2 Z)^n} \Big| \leq \sum_{n=N}^{\infty} \frac{2^n \Gamma[n+(2k/3)]C}{\Gamma(2k/3)\Gamma(n+1)} \Big| \frac{(1+\epsilon)^n s^{(3n/2)}}{(-t^2 Z)^n} \Big|,$$

which converges for $| 2\lambda/t^2 Z | < (1+\epsilon)^{-1}$. Since ϵ is arbitrarily small, the theorem is proved.

5. Inversion formulae. In this section we shall determine the inverse of the various integral operators under consideration, i. e., the analytic func-

tion $f(Z)$ which, when substituted in an operator of the form (1.2), yields the given (real) solution $\psi(Z, \bar{Z}) = \boldsymbol{O}(f)$. The determination of the inverse reduces such questions as the determination of the domain of regularity, the determination of the character of the singularities, etc., to the corresponding problems in the theory of analytic functions of one complex variable.

As we stressed in section **1**, it is sometimes of interest to consider the integral operator of the first kind with the point of reference *different from the origin.* Such an operator may be written in the form [17]

$$\boldsymbol{p}_1(g, Z_0) = R(Z', \bar{Z}')\psi^*(Z', \bar{Z}'), \tag{5.1}$$

where

$$\psi^*(Z', \bar{Z}') = g(Z') + \sum_{n=1}^{\infty} \frac{(2n)!}{2^{2n}n!} P^{(n)}(Z', \bar{Z}') \int_0^{Z'} \cdots \int_0^{Z'_{n-1}} g(Z'_n) dZ_n \cdots dZ'_1, \tag{5.2}$$

and

$$Z' = Z - Z_0, \qquad \bar{Z}' = \bar{Z} - \bar{Z}_0, \tag{5.3}$$

and the functions $P^{(n)}$ satisfy the recurrence relations,

$$P_{\bar{Z}'}^{(1)} + 2F = 0; \qquad (2n+1)P_{\bar{Z}'}^{(n+1)} + 2P_{Z'\bar{Z}'}^{(n)} + 2FP^{(n)} = 0 \quad (n = 1, 2, \cdots), \tag{5.4}$$

and are such that

$$P^{(n)}(Z', 0) = 0 \qquad (n = 1, 2, \cdots). \tag{5.5}$$

Here $F = F[(1/2)(Z' + \bar{Z}') + (1/2)(Z_0 + \bar{Z}_0)]$ is the function introduced in (4.3), where $\lambda = (1/2)(Z + \bar{Z})$, see (1.1a).

THEOREM 5.1. *Let $\psi(Z, \bar{Z})$ be a real solution of* (1.1a) *which is regular in the neighborhood of $Z = Z_0$, $\bar{Z} = \bar{Z}_0$, and is such that*

$$2\psi(Z, \bar{Z}) = \boldsymbol{p}_1(g; Z_0) + \overline{\boldsymbol{p}_1(g; Z_0)}. \tag{5.6}$$

Then

$$2\psi(Z, \bar{Z}_0) = g(Z') + \bar{R}(0, Z')\bar{g}(0), \quad 2\psi(Z_0, \bar{Z}) = R(0, \bar{Z}')g(0) + \bar{g}(\bar{Z}'). \tag{5.7}$$

Proof. This follows immediately from (5.1), (5.2) and (5.5), by the substitution of $\bar{Z}' = 0$ and $Z' = 0$, respectively.

---

[17] (5.1) is somewhat altered form of P(f), see (1.2). In (5.13) the relation between the corresponding f and g is given.

Similarly, we now consider the integral operator of the second kind with an arbitrary reference point Z_0. Such an operator is defined by [18]

$$(5.8) \qquad \mathrm{P}_{21}(f; Z_0) = \int_{-1}^{+1} E_{21}(Z', \bar{Z}', t) f[(Z'/2)(1-t^2)]\, dt/(1-t^2)^{1/2}$$

(see (1.2)), where

$$(5.9) \qquad E_{21}(Z', \bar{Z}'; t) = H(2\lambda')[1 + \sum_{n=1}^{\infty} (t^2 Z')^n Q^{(n)}(2\lambda')],$$

$$2\lambda = Z' + \bar{Z}'; \quad 2\lambda_0 = Z_0 + \bar{Z}_0; \quad H(2\lambda') = \exp[-\int_{-\infty}^{2\lambda'} N(\lambda + \lambda_0) d\lambda],$$

and where the functions $Q^{(n)}(2\lambda')$ are given by the recursion formula

$$(5.10) \qquad (2n+1)[dQ^{(n+1)}/d\lambda'] + [d^2 Q^{(n)}/d\lambda'^2] + 4F(\lambda' + \lambda_0) Q^{(n)} = 0,$$

$$[dQ^{(1)}/d\lambda'] + 4F(\lambda' + \lambda_0') = 0,$$

and are such that

$$(5.11) \qquad Q^{(n)}(\lambda') = 0, \text{ for } \lambda' = 0, \qquad (n = 1, 2, \cdots).$$

Remark. The domain of regularity R of this operator is given by $(\lambda - \lambda_0)^2 + \theta^2 < 4\lambda^2$. This may be established by the method of [7], pp. 36-39, replacing λ by $\lambda' = \lambda - \lambda_0$ throughout. The operator (5.8) may be written in the alternate form

$$(5.12) \qquad \boldsymbol{p}_{21}(g; Z_0) = H(2\lambda')\psi^*(Z', \bar{Z}'),$$

where $\psi^*(Z'\bar{Z}')$ has the same form as that given in (5.2), except that $P^{(n)}(Z', \bar{Z}')$ are to be replaced by $Q^{(n)}(2\lambda')$. Furthermore, f and g are related by

$$(5.13) \qquad g(Z') = \int_{-1}^{+1} f[(Z'/2)(1-t^2)]\, dt/(1-t^2)^{(1/2)},$$

THEOREM 5.2. *Let there be given a (real) solution*

$$\psi(Z', \bar{Z}') = \Psi(\lambda', \theta') = H(2\lambda')\psi^*(Z', \bar{Z}') = H(2\lambda')\Psi^*(\lambda', \theta')$$
$$= (1/2)(\boldsymbol{p}_{21}(g; Z_0) + \overline{\boldsymbol{p}_{21}(g; Z_0)}),$$

where

$$g(Z) = \sum_{n=0}^{\infty} \mu_n Z^n, \qquad \mu_n = \mu_{n1} + i\mu_{n2}.$$

[18] The operator P_{21} with the reference point $Z_0 = 0$ has been considered in section 4.

Furthermore, let

$$\Psi^*(0,\theta') = \sum_{n=0}^{\infty} \beta_n \theta'^n, \tag{5.14}$$

$$\Psi^*_{\lambda'}(0,\theta') = [H(0)]^{-1}\Psi_{\lambda'}(0,\theta') - H_{\lambda'}(0)[H(0)]^{-2}\Psi(0,\theta') = \sum_{n=0}^{\infty} \gamma_n \theta'^n.$$

Then μ_n are related to β_n and γ_n as follows:

$$\mu_{2n1} = (-1)^n \beta_{2n}, \quad \mu_{2n+1\,2} = (-1)^{n+1}\beta_{2n+1} \qquad (n = 0, 1, 2, \cdots), \tag{5.15}$$

$$\mu_{2n+2-k\,k} = \begin{vmatrix} D_0^{(k)} & D_1^{(k)} & \cdots & D_{n-1}^{(k)} & D_n^{(k)} \\ a_{00}^{(k)} & a_{10}^{(k)} & \cdots & a_{n-1\,0}^{(k)} & a_{n0}^{(k)} \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & \cdots & a_{0\,n-1}^{(k)} & a_{n\,n-1}^{(k)} \end{vmatrix} \qquad \begin{matrix} (n = 0, 1, 2, \cdots; \\ k = 1, 2), \end{matrix} \tag{5.16}$$

where $a_{0\nu} = -1$, $\nu = 0, \cdots, n-1$, and where

$$D_n^{(1)} = (-1)^n(2n+1)^{-1}\gamma_{2n} - [(2n+1)!]^{-1}\sum_{j=1}^{n}(-1)^{n-j}C_{2j}A_{2j}[(2n-2j)!]\beta_{2n-2j}, \tag{5.17a}$$

$$a_{nm}^{(1)} = -[(2n+1)!]^{-1}C_{2n-2m-1}A_{2n-2m-1}(2m+1)!, \tag{5.17b}$$

$$C_n = (2n)!\,2^{-2n}(n!)^{-1}, \qquad A_n = Q^{(n)}(\lambda')\Big|_{\lambda=0}, \tag{5.17c}$$

and

$$D_n^{(2)} = (-1)^n 2^{-1} n^{-1}\gamma_{2n-1} - [(2n)!]^{-1}\sum_{j=1}^{n-1} C_{2j}A_{2j}[(2n-2j-1)!](-1)^{n-j}\beta_{2n-2j-1}, \tag{5.18a}$$

$$a_{nm}^{(2)} = -[(2n)!]^{-1}C_{2n-m-1}A_{2n-m-1}(2m)!. \tag{5.18b}$$

Proof.

$$2\Psi^*(\lambda',\theta') = \sum_{n=0}^{\infty}\mu_n(\lambda'+i\theta')^n + C_1Q^{(1)}(\lambda')\sum_{n=0}^{\infty}(n+1)^{-1}\mu_n(\lambda'+i\theta')^{n+1}+\cdots \tag{5.19}$$

$$+\sum_{n=0}^{\infty}\bar{\mu}_n(\lambda'-i\theta')^n + C_1Q^{(1)}(\lambda')\sum_{n=0}^{\infty}(n+1)^{-1}\bar{\mu}_n(\lambda'-i\theta')^{n+1}+\cdots,$$

and therefore

$$2\Psi^*(0,\theta') = \sum_{n=0}^{\infty}(\mu_{n1}[(i)^n + (-i)^n] + \mu_{n2}[(i)^{n+1} + (-i)^{n+1}]\theta'^n = 2\sum_{n=0}^{\infty}\beta_n\theta'^n. \tag{5.20}$$

Equations (5.15) follow immediately from this. Furthermore,

$$(5.21)\quad 2[d\Psi^*(\lambda',\theta')/d\lambda'] = \sum_{n=0}^{\infty} n\mu_n(\lambda'+i\theta')^{n-1} + C_1Q^{(1)}(\lambda')\sum_{n=0}^{\infty}\mu_n(\lambda'+i\theta')^n$$
$$+ C_1[dQ^{(1)}(\lambda')/d\lambda']\sum_{n=0}^{\infty}(n+1)^{-1}\mu_n(\lambda'+i\theta')^{n+1}+\cdots$$
$$+\sum_{n=0}^{\infty} n\bar{\mu}_n(\lambda'-i\theta')^{n-1} + C_1Q^{(1)}(\lambda')\sum_{n=0}^{\infty}\bar{\mu}_n(\lambda'-i\theta')^n$$
$$+ C_1[dQ^{(1)}(\lambda')/d\lambda']\sum_{n=0}^{\infty}(n+1)^{-1}\bar{\mu}_n(\lambda'-i\theta')^{n+1}+\cdots$$

Whence, by (5.11) and (5.17d),

$$(5.22)\quad 2[d\Psi^*(\lambda',\theta')/d\lambda']\Big|_{\lambda'=0} = \sum_{n=0}^{\infty} n\mu_n(i)^{n-1}\theta'^{n-1}$$
$$+ C_1A_1\sum_{n=0}^{\infty}(n+1)^{-1}(i)^{n+1}\mu_n\theta'^{n+1}$$
$$+ C_2A_2\sum_{n=0}^{\infty}(n+1)^{-1}(n+2)^{-1}(i)^{n+2}\mu_n\theta'^{n+2}+\cdots$$
$$+\sum_{n=0}^{\infty} n\bar{\mu}_n(-i)^{n-1}\theta'^{n-1} + C_1A_1\sum_{n=0}^{\infty}(n+1)^{-1}(-i)^{n+1}\bar{\mu}_n\theta'^{n+1}$$
$$+ C_2A_2\sum_{n=0}^{\infty}(n+1)^{-1}(n+2)^{-1}(-i)^{n+2}\bar{\mu}_n\theta'^{n+2}+\cdots$$
$$= \sum_{n=0}^{\infty}\gamma_n\theta'^n.$$

From this it follows that

$$(5.23)\quad \mathrm{Re}\{i^n[(n+1)\mu_{n+1} + C_1A_1n^{-1}\mu_{n-1} + C_2A_2[n(n-1)]^{-1}\mu_{n-2}$$
$$+\cdots+C_nA_n\mu_0(n!)^{-1}]$$
$$=\gamma_n \qquad (n=0,1,\cdots).$$

It follows from this that

$$(5.24)\quad \gamma_{2n}(-1)^n = (2n+1)\mu_{2n+1\,1} + [(2n)!]^{-1}\sum_{j=1}^{2n}(2n-j)!C_jA_j\mu_{2n-j\,1}$$

$$(5.25)\quad \gamma_{2n+1}(-1)^{n+1} \qquad (n=0,1,\cdots)$$
$$= (2n+2)\mu_{2n+2\,2} + [(2n+1)!]^{-1}\sum_{j=1}^{2n+1}(2n+1-j)!C_jA_j\mu_{2n+1-j\,2}$$
$$(n=0,1,\cdots).$$

From the first of these relations and (5. 15), we obtain

$$\mu_{2n+1\,1} = \sum_{j=0}^{n-1} a_{nj}^{(1)} \mu_{2j+1\,1} + D_n^{(1)} \qquad (n = 0, 1, \cdots), \tag{5.26}$$

while from the second, we obtain

$$\mu_{2n2} = \sum_{j=0}^{n-1} a_{nj}^{(2)} \mu_{2j} + D_n^{(2)} \qquad (n = 0, 1, \cdots). \tag{5.27}$$

The expressions (5. 16) now follow.

Theorem 5. 2 gives us the coefficients of the associate of the second kind in terms of the initial values χ_1, χ_2. The connection between the coefficients β_n, γ_n on one side, and μ_n on the other, can be interpreted as a projection in the coefficient space. While the associate of the integral operator of the first kind of $\psi = \Sigma A_{mn} z^m \bar{z}^n$, $A_{nm} = \bar{A}_{mn}$, can be determined in terms of the A_{n0}'s *independently* of the equation, the above-mentioned projection depends upon the equation, since the elements of the determinant (5. 16) depend on the $Q^{(n)}$'s. (See (5. 17b), (5. 17c)).

Our result represents much more than the solution of the initial value problem. All properties of the associate g (a function of a complex variable), or f (see (5. 13)), can be "translated" into theorems on the behavior of ψ in the domain R (cf. the remark on p. 461) of the regularity of the operator. For details, see [6], p. 299. As an example, it is of interest to formulate the results which follow from the Hadamard theory of coefficients of entire or meromorphic functions. If the coefficients of f satisfy the well-known Hadamard conditions, so that f has in a domain $B \subset R$, poles as its only singularities, then ψ will, at the corresponding points, have certain branch points of order $n + (1/2)$, n integer. If $f = f_1/f_2$, where f_1 and f_2 are functions which are regular in B, the representation

$$\psi(Z', \bar{Z}') = \int_{t=-1}^{1} E_{21}(Z', \bar{Z}', t) f(Z'(1 - t^2)/2) dt/(1 - t^2)^{1/2}, \tag{5.28}$$

holds in B.

Remark. ψ is an infinitely many-valued function possessing branch point at the point p_ν where f_2 has zeros, and we have to specify what branch of it is represented by (5. 28).

If on the other hand g satisfies the Hadamard conditions, so that we can write $g = (g_1/g_2)$, where g_1 and g_2 are regular in B, then we can write

$$(5.29)\quad \psi(Z', \bar{Z}') = -(2\pi)^{-1} \int_{t=-1}^{1} E_{21}(Z', \bar{Z}', t)$$

$$\times \int_{\tau=-1}^{1} g_1[Z'(1-\tau^2)(1-t^2)/2]/g_2[Z'(1-\tau^2)(1-t^2)/2](1-t^2)^{-1/2}\tau^{-2}d\tau dt$$

or [19]

$$\psi(Z', \bar{Z}') = H(\lambda')[(g_1/g_2) + \sum_{n=1}^{\infty} [2^n(n!)^2]^{-1}(2n)!Q^{(n)}(2\lambda') \int_0^{Z'} (Z'-\zeta)^{n-1}(g_1(\zeta)/g_2(\zeta))d\zeta].$$

The above formulas can be used for the study of the growth and value distribution in the neighborhood of a point of accumulation of singularities of ψ. Thus, the theory of integral operators gives us insight into the structure of the singularities of a solution of a differential equation.

In this connection it will be of interest to clarify the possibility of the representation of *single-valued* solutions of (1.1) in the neighborhood of a point of accumulation of single-valued singularities.

It is well known that the only singularities of linear differential equations of elliptic type with analytic coefficients which are single-valued in the real plane are fundamental solutions, say $\psi^{(L1)}(Z, Z_0) = A(\log\zeta + \log\bar{\zeta}) + B$, $A \equiv A(Z, \bar{Z}; Z_0, \bar{Z}_0)$, $B \equiv B(Z, \bar{Z}; Z_0, \bar{Z}_0)$, $\zeta = Z - Z_0$, $\bar{\zeta} = \bar{Z} - \bar{Z}_0$, and their derivatives with respect to the parameters Z_0, $\bar{Z}_0$. (See [8], p. 473). Let Z_ν be an infinite set of points in a domain B. The question arises whether there exists an infinite sum

$$\phi = \sum_{\nu=1}^{\infty} [\psi^{(L1)}(Z, Z_\nu) + H_\nu(Z)],$$

of such singularities with conveniently chosen solutions H_ν's, which are regular in B and possess the property that $D(\phi) = \int\int_B \phi^2 w\, dx\, dy < \infty$. Here w is some (positive) weight function. In order to answer this question, we can repeat the considerations carried out in [2], p. 143-147, where we introduced fundamental solutions which are normalized with respect to the domain B. (They possess the property that they are orthogonal to every regular solution ϕ with finite $D(\phi)$ of the differential equation). The considerations of [2] can be repeated in this case, and we get necessary and sufficient conditions for the existence of functions ϕ, $D(\phi) < \infty$, possessing the above singularities at the prescribed points Z_ν. If sufficient conditions are satisfied, the above considerations will also give us a representation of the function ϕ in terms

[19] If g_2 is a polynomial, the above integral can be evaluated in terms of elementary (i.e., algebraic and logarithmic) functions.

of the kernel function of a complete orthonormal system with a weight function w.

6. Other representations of the integral operator of the second kind. In the simplified case, in addition to the representation (3.8) for P_{21} and (3.9) for P_{22}, we can obtain two further representations of (3.7) in powers of $v = 1 - u$ and $(u-1)/u$. We shall show that similar developments hold in the general case.

The function [20] $E^* = E/H$, considered as a function of

$$u = t^2 Z/2\lambda, \quad \text{and} \quad s = (-\lambda)^{2/3}, \tag{6.1}$$

satisfies the differential equation

$$\begin{aligned} u(1-u)E^*_{uu} + (1/2 - 2u)E^*_u - (5/36)E^* + (2s/3)(2u-1)E^*_{us} \\ + 9^{-1}s(2+3u^{-1})E^*_s - (4/9)s^2E^*_{ss} - 4(S(s) - (5/144))E^* = 0. \end{aligned} \tag{6.2}$$

If we now introduce

$$\tilde{E}^* = E^* u^{-1/2}, \tag{6.3}$$

we find that $\tilde{E}^*$ satisfies the differential equation

$$\begin{aligned} u(1-u)\tilde{E}^*_{uu} + ((3/2) - 3u)\tilde{E}^*_u - [4S(s) + (8/9)]\tilde{E}^* \\ + (8/9)s\tilde{E}^*_s + (2s/3)(2u-1)\tilde{E}^*_{us} - (4s^2/9)\tilde{E}^*_{ss} = 0. \end{aligned} \tag{6.4}$$

and upon introducing the variable

$$v = 1 - u = 1 - t^2 Z(2\lambda)^{-1}, \tag{6.5}$$

we have

$$\begin{aligned} v(1-v)\tilde{E}^*_{vv} + [(3/2) - 3v]E^*_v - [4S(s) + (8/9)]\tilde{E}^*_s \\ + (8s/9)\tilde{E}^*_s + (2s/3)(2v-1)\tilde{E}^*_{vs} - (4s^2/9)\tilde{E}^*_{ss} = 0. \end{aligned} \tag{6.6}$$

It will be observed that equations (6.4) and (6.6) are identical, except that the role of u and v is interchanged. It therefore follows that (6.6) has solutions of the form

$$\tilde{E}^* = v^{-1/2}[(-\lambda)^{(k/2)-(1/2)} + \sum_{n=1}^{\infty} (-2\lambda v)^n Q^{(nk)}(2\lambda) \qquad (k = 1, 2), \tag{6.7}$$

and

$$\tilde{E}^* = v^{-1/2} \sum_{n=0}^{\infty} [[q^{(nk)}(\lambda)](-2\lambda v)^{-[n-(1/2)+(2k/3)]}] \qquad (k = 1, 2), \tag{6.8}$$

whence, in the original variables,

$$E^* = tZ^{1/2}(2\lambda - t^2 Z)^{-1/2}[(-\lambda)^{(k/2)-(1/2)} + \sum_{n=1}^{\infty} (t^2 Z - 2\lambda)^n Q^{(n)}(2\lambda)] \qquad (k = 1, 2), \tag{6.9}$$

[20] In this section we consider both E_{21} and E_{22}, but the subscripts have been omitted.

valid for $|(t^2Z-2\lambda)/2\lambda|<1$, and

$$(6.10)\qquad E^*=tZ^{1/2}[(2\lambda-t^2Z)^{-1/2}\sum_{n=0}^{\infty}[q^{(nk)}(\lambda)](t^2Z-2\lambda)^{-[n-(1/2)+(2k/3)]}]\qquad(k=1,2),$$

valid for $|(t^2Z-2\lambda)/2\lambda|>1$.

7. On solutions of the initial value problem in the large. One of our aims is to consider the problem of the solution of the initial value problem in the large. More precisely: if on a segment of the transition line, the value of a solution and of its derivative in a direction normal to the line are given, how far can the solution be continued into the elliptic region, what are the locations and the character of its singularities, etc.? In the present section, we shall obtain a partial answer to these questions. The properties thus obtained are, to a certain extent, independent of the function $F(\lambda)$ (or N). We shall now obtain a sufficient condition which will guarantee the existence of a solution $\psi(\lambda,\theta)$ of equation (3.1) which is regular in a subregion B of the elliptic region, and which satisfies given initial data on the transition line $\lambda=0$. The problem is reduced to the determination of the domain of regularity of an analytic function of a complex variable in terms of the coefficients of its Taylor development. This question can be answered by a number of results in the theory of analytic functions, and we shall not enter into it here.

Let Γ be a simple curve which connects the points -1 and $+1$ of the *complex* t-plane and lies completely in $1\leq|t|\leq A<\infty$, A sufficiently large. For each point ζ, we denote by $c(\zeta;\Gamma)$ the curve traced out by the values $(1/2)\zeta(1-t^2)$ as t traverses Γ from -1 to $+1$. Furthermore, we denote by $T_\Gamma(B)$ the domain

$$(7.1)\qquad T_\Gamma(B)=\sum_{\zeta\varepsilon B}c(\zeta;\Gamma).$$

If, for instance, Γ is the upper half of the unit circle, then $c(\zeta;\Gamma)$ is the circle $|z-\zeta/2|=|\zeta/2|$, so that $T_\Gamma(B)$ consists of the union of all circles passing through the origin and points $\mathfrak{P}\varepsilon B$, and having the segment $\widehat{0\mathfrak{P}}$ as diameter. More generally, as t traverses an arbitrary curve from -1 to $+1$, the expression $(\zeta/2)(1-t^2)$ varies on a closed curve c which passes through the origin. Thus $T_\Gamma(B)$ consists of the union of closed curves $c(\zeta,\Gamma)$, each passing through the origin, and each obtained from the fixed curve c by a dilation of amount $|\zeta|$ and a rotation of amount $\arg\zeta$. It is clear that any closed curve passing through the origin may be selected as the fixed curve c, and that the origin will be a boundary point for all domains $T_\Gamma(B)$.

THEOREM 7.1. *Let $\chi_1(\theta)$ and $\chi_2(\theta)$ be two functions of θ which possess representations of the form*

$$\chi_j(\theta) = \sum_{\nu=0}^{\infty} a_\nu^{(j)} \theta^\nu, \qquad 0 \leq \theta \leq \theta_1, \theta_1 > 0, (j = 1, 2). \tag{7.2}$$

Let a function $f(\zeta)$ be defined by

$$f(\zeta) = \zeta^{1/6} \sum_{\nu=0}^{\infty} c_\nu \zeta^\nu, \tag{7.3}$$

where

$$c_\nu = -(-2i)^{1/6}[3^{1/2}\pi S_0^2 \mathrm{Im}(A_2\bar{A}_1)]^{-1} \times [-a_\nu^{(1)}\bar{d}_0 J_\nu^{(1)} + \sum_{k=1}^{2}(-1)^k a_\nu^{(k)}\bar{d}_k J_\nu^{(2)}], \tag{7.4}$$

with $\mathrm{Im}[A_2\bar{A}_1] \neq 0$, *and* [21]

$$d_0 = -(2/3)i^{3/2}S_0A_2, \quad d_1 = -(2^{5/3}/3)i^{1/6}S_0S_1A_1 \quad d_2 = -i^{1/6}S_0A_1. \tag{7.5}$$

$$J_\nu^{(1)} = (1/2)e^{-(2/3)n\pi i}(e^{-(2/3)\pi i} - 1)[\Gamma(-1/3)\Gamma(\nu+1)/\Gamma(\nu+2/3)] \quad (n = 0, \pm 1), \tag{7.6}$$

$$J_\nu^{(2)} = (1/2)e^{-(4/3)n\pi i}(e^{-(4/3)\pi i} - 1)[\Gamma(1/3)\Gamma(\nu+1)/\Gamma(\nu+4/3)] \quad (n = 0, \pm 1).$$

If a domain B situated in $\mathbf{E}[3^{1/2} \mid \lambda \mid < \theta, \theta > 0, \lambda \leq 0]$, contains the segment $0 \leq \theta \leq \theta_1$, and if $f(\zeta)$ is regular in $T_{\mathrm{I}}(B)$, then

$$\psi(\lambda, \theta) = \mathrm{Im}[\int_{\mathrm{I}} E(Z, \bar{Z}, t) f[Z(1-t^2)/2](1-t^2))^{-1/2}\, dt], \tag{7.7}$$

where $E = A_1E^{(1)} + [Z(1-t^2)/2]^{2/3}A_2E^{(2)}$, is a solution of the differential equation (2.3), which is regular in B and is such that [22]

$$\lim_{\lambda \to 0^-} \psi(\lambda, \theta) = \chi_1(\theta), \tag{7.8a}$$

$$\lim_{\lambda \to 0^-} (-\lambda)^{1/3}[\partial\psi(\lambda, \theta)/\partial\lambda] = \chi_2(\theta). \tag{7.8b}$$

Proof. With t on I and $Z \in B$, by Theorem 4.1, $E(Z, \bar{Z}, t)$ is regular. By assumption, $f[Z(1-t^2)/2]$ is also regular there. Hence, (7.7) is a solution of the differential equation (3.1). The boundary relations (7.8a) and (7.98) follow [22] from Theorem 6.1 of [9].

---

[21] The three independent possibilities in (7.6), are covered by $n = 0$, $n = 1$, $n = -1$, respectively.

[22] It follows from (3.3) that $\lim_{\lambda \to 0^-} (-\lambda)^{-1/3} d\lambda/d\mathrm{H} = (3a_1/2)^{1/3}$, so that the limit in (7.8b) essentially coincides with $\lim_{\mathrm{H} \to 0} \psi(\mathrm{H}, \theta)$.

It should be further noted that the initial value problem for the equation $\psi_{\mathrm{HH}} + (-\mathrm{H})^s\psi_{\theta\theta} = 0$, $s > -1$, has been studied in the papers [10], [15], [18], and formulas for the representation of the solution have been obtained for these particular cases.

THEOREM 7.2. *Let $\chi_1(\theta)$ and $\chi_2(\theta)$ be two real functions of θ which are regular in $0 \leq \theta \leq \theta_1$, and which possess expansions of the form (7.2). If the function $g(\zeta) = \zeta^{-1/6} f(\zeta) = \sum_{\nu=0}^{\infty} c_\nu \zeta^\nu$, where c_ν are the linear combinations of the $a_\nu^{(j)}$ given by (7.4), possesses at the point $\zeta = \alpha$, $(4/3)\pi < \arg \alpha < 2\pi$, an expansion of the form*

$$g(\zeta) = \zeta + \sum_{\nu=0}^{\infty} a_\nu (\zeta - \alpha)^{-\nu}, \tag{7.9}$$

valid for $0 < |\zeta - \alpha| < \infty$, then there exists a solution $\psi(\lambda, \theta)$ of the differential equation (3.1) which is regular in $\mathrm{E}[3^{1/2}|\lambda| < |\theta|, \theta > 0, |\lambda| \leq 0]$, and such that

$$\lim_{\lambda \to 0^-} \psi(\lambda, \theta) = \chi_1(\theta), \tag{7.10a}$$

$$\lim_{\lambda \to 0^-} (-\lambda)^{1/3} \psi_\lambda(\lambda, \theta) = \chi_2(\theta), \tag{7.10b}$$

and which possesses a singularity at the point of infinity.

Proof. We consider the function

$$\psi(\lambda, \theta) = \mathrm{Im}\Big[\int_{\mathrm{I}} E(Z, \bar{Z}, t) f[Z(1-t^2)/2]\, dt/(1-t^2)^{1/2}\Big] = \mathrm{Im}[\mathrm{P}_{22}(f)], \tag{7.11}$$

where I is the upper half of the unit circle, and

$$E(Z, \bar{Z}, t) = \{\exp[-\int_{-\infty}^{2\lambda} 2N(t)dt]\}[A_1 E^{*(1)} + [Z(1-t^2)/2]^{2/3} A_2 E^{*(2)}]. \tag{7.12}$$

For $Z \in R \equiv \mathrm{E}[3^{1/2}|\lambda| < |\theta|, \theta > 0, \lambda \leq 0]$ and $t \in \mathrm{I}$, it is clear that the points $\zeta = Z(1-t^2)/2$ will be confined to the sector $0 \leq \arg \zeta \leq (4/3)\pi$. Hence if α is confined to the sector $(4/3)\pi \leq \arg \zeta < 2\pi$, and $g(\zeta)$ possess the expansion (7.9), then $f[Z(1-t^2)/2]$ will be regular for $Z \in R$, $t \in \mathrm{I} - 0$ (the endpoints, $t = \pm 1$ give no trouble in the integration). Thus, by the preceding theorem, $\psi(\lambda, \theta)$ is a solution of (3.1) which is regular in $\mathrm{E}[3^{1/2}|\lambda| < \theta, 0 < \theta < \infty, \lambda \leq 0]$, and is such that the initial value relations (4.10) are valid.

Now $f(\zeta) = \zeta^{7/6} + \zeta^{1/6} \sum_{\nu=0}^{\infty} a_\nu (\zeta - \alpha)^{-\nu} = \zeta^{7/6} + \zeta^{1/6} h(\zeta)$, where $h(\zeta) = O(1)$ for $\zeta \to \infty$.

Also $\psi(\lambda, \theta) = \mathrm{Im}[\mathrm{P}_{22}(f)] = \mathrm{Im}[\mathrm{P}_{22}(\zeta^{7/6})] + \mathrm{Im}[\mathrm{P}_{22}(\zeta^{1/6} h(\zeta))]$. We shall show that on the line $\lambda = 0$ the first integral becomes infinite, while the second remains bounded. From (7.5), (7.6) and (7.7), it is clear that on $\lambda = 0$, $E(Z, \bar{Z}, t)$ becomes $E(i\theta, -i\theta, t) = \mathrm{const}\,(-it^2\theta)^{-1/6}$. Thus

$P_{22}(\zeta^{7/6})\big|_{\lambda=0} = \text{const}(i\theta/2)\int_I [(1-t^2)^{1/2}(-t)^{-1/3}]dt$. Depending upon the determination of the root, we have [23]

$$\int_I [(1-t^2)^{1/2}(-t)^{-1/3}]dt = (-1)^{2/3}e^{-(4/3)n\pi i}(e^{-(4/3)\pi i}-1)$$
$$[\Gamma(1/3)\Gamma(5/6+4/6)][\Gamma(1+5/6)]^{-1} \quad (n=0,\pm 1)$$

(see [9], p. 889). Thus $\text{Im}[P_{22}(\zeta^{7/6})]\big|_{\lambda=0} = e_1\theta$, $e_1 \neq 0$, so that

$$\lim_{\theta\to\infty}[\text{Im}\, P_{22}(\zeta^{7/6})]\big|_{\lambda=0} = \infty.$$

Furthermore,

$$\text{Im}[P_{22}(\zeta^{1/6}h(\zeta))]\big|_{\lambda=0} = \text{Im}\int_I (-t)^{-1/3}(1-t^2)^{-1/2}h(i\theta(1-t^2)/2)\,dt.$$

Thus

$$|\,\text{Im}[P_{22}(\zeta^{1/6}h(\zeta))]\big|_{\lambda=0} \leq \int_I |(-t)^{-1/3}(1-t^2)^{-1/2}h(i\theta(1-t^2)/2)|\,|\,dt\,| = O(1),$$
for $\theta\to\infty$.

In (2.1) and (2.2) of section 2, we have considered a set of particular solutions ψ_{nk} (introduced in [4]). In [7], the operator R_2 was introduced [24] transforming the analytic function $\sum_{n=0}^{\infty}(\alpha_n+i\beta_n)\zeta^n$, $\zeta=\theta+i\text{H}$ as follows:

(7.13) $\quad R_2[\Sigma(\alpha_n+i\beta_n)\zeta^n] = \Sigma[\alpha_n(\theta+i\text{H})^{[n]} + i\odot\beta_n(\theta+i\text{H})\{n\}]$.

Here $\text{Im}(\theta+i\text{H})^{[n]} = \psi_{n1}$, $\text{Im}(\theta+i\text{H})\{n\} = \psi_{n2}$. For the symbol $\odot$ see [8], p. 455. It is of interest that the operator $\text{Im}\,R_2$ coincides in certain cases with the integral operator $\text{Re}\,P_{22}$ considered in section 4.

THEOREM 7.3. *Let $\chi_1(\zeta)$ and $\chi_2(\zeta)$ be regular for $|\zeta|<\zeta_0$, and let $l(\text{H})$ be regular at $\text{H}=0$. Then*

(7.14) $\quad \text{Im}\,R_2[\int_0^{\zeta}\chi_2(t)dt + i\chi_1(\zeta)]$
$$= \text{Re}\int_{t=-1}^{+1} E_{22}(\lambda,\theta,t)f(Z(1-t^2)/2))\,dt/(1-t^2)^{1/2},$$

where $f(\zeta)$ is given by

(7.15) $\quad f(\zeta) = -\{(-2i\zeta)^{1/6}/3^{1/2}\pi S_0^2\,\text{Im}[A_2\bar{A}_1]\}$
$$\times[-\bar{d}_0\int_{C_2} t^{-1/3}\chi_1(\sigma)\,dt - \bar{d}_1\int_{C_2} t^{-5/3}\chi_1(\sigma)\,dt$$
$$+ (2/3a_1)^{1/3}\bar{d}_2\int_{C_2} t^{-5/3}\chi_2(\sigma)\,dt], \qquad \sigma = -2i\zeta(1-t^2),$$

and where the constants appearing here have been defined in (7.5).

[23] See footnote 21.

[24] In [7] the operator R_2 was designated by P_2.

Proof. It is easily shown that

$$(7.16)\quad \operatorname{Im} R_2\Big[\int_0^{\zeta} \chi_2(t)\,dt + i\chi_1(\zeta)\Big] = \sum_{n=1}^{\infty} (1/n)\,a_{n-1}^{(2)}\psi_{n1}(H,\theta) + \sum_{n=0}^{\infty} a_n^{(1)}\psi_{n2}(H,\theta) = \psi(H,\theta),$$

where $\chi_k(\theta) = \sum_{\nu=0}^{\infty} a^{(k)}\theta^{\nu}$ $(k = 1, 2)$.

If we assume that $l(H)$ is an analytic function of H, it follows from Theorem 7.1 that $\psi(H,\theta)$ is a solution of the differential equation (1.3) which is analytic in H and θ in a sufficiently small neighborhood of, say, $N = [-\theta_1 \leq \theta \leq \theta_1, H_0 \leq H \leq 0]$, $\theta_1 > 0$, and is such that $\psi(0,\theta) = \chi_1(\theta)$, $\partial\psi(H,\theta)/\partial H\big|_{H=0} = \chi_2(\theta)$. By Theorem 6.1 of [9], the expression $\operatorname{Re}\int_{-1}^{1} E_{22}(\lambda,\theta;t) f(Z(1-t^2)/2)(1-t^2)^{-1/2}dt$ is also a solution, possessing the initial values $\psi(0,\theta) = \chi_1(\theta)$ and $\partial\psi(H,\theta)/\partial H)\big|_{H=0} = \chi_2(\theta)$ on the line $H = 0$. We shall now show that this solution must also be regular in the neighborhood N. From this it will follow by the uniqueness theorem (cf. the remark on p. 449) that the two solutions considered must be identical.

It is clear, first of all, that the function

$$\int_{-1}^{+1} E_{22}(\lambda,\theta;t) f(Z(1-t^2)/2)(1-t^2)^{-1/2}dt$$

must be regular in $s^{1/2}$, where $s = (-\lambda)^{2/3}$ (see (5.7) of [9]). Since by (3.3), H is a regular function of s in the vicinity of $s = 0$, it suffices to show that this integral must be regular in s.

If we replace in the differential equation (1.1a) the variable λ by s, we obtain

$$(7.17)\quad 4s^{-1}(\partial^2\psi/\partial s^2) + 9(\partial^2\psi/\partial\theta^2) = 24s^{-1}[\beta_1 + \beta_2 s + \cdots](\partial\psi/\partial s) = 0.$$

Since the above solution is known to be regular in $s^{1/2}$, we have

$$(7.18)\quad \psi = \sum_{n=0}^{\infty} a_n s^n + \gamma_p s^{(p+1)/2} + \gamma_{p+1} s^{(p+3)/2} + \cdots,\ p \text{ even},\ \geq 2.$$

In view of the fact that the fractional powers do not combine with the integer powers, we need merely substitute the second portion of (7.16) into (7.17). When this is done, and when we compare the coefficient of $s^{(p-5)/2}$, we obtain $4((p+1)/2)((p-1)/2)\gamma_p = 0$. Hence $\gamma_p = 0$, and this shows (by induction) that fractional powers cannot appear in a solution of the form (7.18).

8. Connection between the results obtained by the operator method and the general theory of functions of two complex variables. It is of considerable interest that the results obtained by the operator method can be considered as complementing an important portion of the general theory of analytic functions of several complex variables. While in the case of functions of one complex variable, the methods of Nevanlinna, Ahlfors, and others establish relations between the growth of the function and value distribution in the whole plane, generalizations of these methods (by exploiting the idea of the distinguished boundary surface and that of the extended class) lead us to relations between the growth and value distribution in certain (non-analytic) *three*-dimensional manifolds of the (four diemnsional) space of two complex variables. These investigations in two variables can be complemented by results establishing relations between coefficients of the series development, say of an entire function, and its growth, so that from properties of the coefficients we can draw conclusions about the growth and the value distribution in certain three-dimensional manifolds of the z_1, z_2-space.

As it has been mentioned in [2], one of the aims of the theory of functions of two complex variables is to investigate the relations between the properties of coefficients, growth and value distribution in certain non-analytic *two*-dimensional manifolds, e. g., in the real plane $y_1 = 0$, $y_2 = 0$.

The fact that a function $f = \sum_{m,n} a_{mn} z^m \bar{z}^n$ of two complex [25] variables $z, \bar{z}$ satisfies a linear differential equation with *entire* coefficients,

$$A = \sum_{\nu,\mu} A_{\nu\mu} z^\nu \bar{z}^\mu, \qquad B = \sum_{\nu,\mu} B_{\nu\mu} z^\nu \bar{z}^\mu, \qquad F = \sum_{\nu,\mu} C_{\nu\mu} z^\nu \bar{z}^\mu,$$

can be expressed by requiring that the a_{mn} satisfy infinitely many relations

$$(8.1)\quad (m+n)(n+1)a_{m+1\,n+1} + \sum_{\nu=1}^{m}\sum_{\mu=0}^{n} A_{\nu\mu}(m-\nu+1)a_{m-\nu+1\,n-\mu} + \sum_{\nu=0}^{m}\sum_{\mu=0}^{n}(n-\mu-1)B_{\nu\mu}a_{m-\nu,\,n-\mu+1} + \sum_{\nu=0}^{m}\sum_{\mu=0}^{n} C_{\nu\mu}A_{m-\nu\,n-\mu} = 0$$
$$(m = 0, 1, 2, \cdots; n = 0, 1, 2, \cdots),$$

where $A_{\nu\mu}$, $B_{\nu\mu}$, $C_{\nu\mu}$ go to zero in a certain way as $\nu \to \infty$, $\mu \to \infty$.

On the other hand, the class of functions which we obtain in this manner has a number of interesting properties which sharply distinguish it from other analytic functions of two complex variables. The method of integral operators indicates that this class of functions has very many properties in common with

---

[25] We note that when $f(z, \bar{z})$ is continued to complex values of x and y, z and $\bar{z}$ are two *independent* variables.

the class of functions $\psi(z, \bar{z}) = f_1(z) + f_2(\bar{z})$, i. e., sums of two functions of *one* variable. For instance, these functions can have singularities only on the surfaces $z = \text{const.}$, $\bar{z} = \text{const.}$ If, on the other hand, our function satisfies a differential equation with *singular* coefficients, say (3. 1), with N satisfying the conditions indicated in section 3, then the character of these solutions, when continued to complex values of the arguments, changes considerably.[26] Thus these functions of two complex variables, in addition to the surfaces $z = \text{const.}$, $\bar{z} = \text{const.}$, can have singularities on the singularity surface of the coefficients, in our case on $z + \bar{z} = 0$.

These results can be generalized to the study of *pairs* of functions. If we consider a pair of solutions ψ_1, ψ_2 of the differential equation (1. 1) (or (3. 1)) in the real space, it is natural to pair the two real solutions whose series developments have the form

$$\psi_1 = \sum_{n=0}^{\infty} \rho^n (a_n \cos n\phi - b_n \sin n\phi) + \sum_{\substack{m=0 \\ m \neq n}}^{\infty} \sum_{n=0}^{\infty} \rho^m (c_{mn}^{(1)} \cos n\phi + d_{mn}^{(1)} \sin n\phi),$$

$$\psi_2 = \sum_{n=0}^{\infty} \rho^n (b_n \cos n\phi + a_n \sin n\phi) + \sum_{\substack{m=0 \\ m \neq n}} \sum_{n=0} \rho^m (c_{mn}^{(2)} \cos n\phi + d_{mn}^{(2)} \sin n\phi),$$

$$z = \rho e^{i\phi}, \qquad \bar{z} = \rho e^{-i\phi}.$$

(We note that the coefficients a_n and b_n can be given arbitrarily, while the coefficients $c_{mn}^{(k)}$, $d_{mn}^{(k)}$ $(k = 1, 2)$, are uniquely determined by the requirements that ψ_k satisfies th equation (1. 1) or the relation (8. 1).

The theory of integral operators has been applied in some cases ([6], pp. 321-330, and [14]) to obtain theorems on value distribution in the real plane $y_1 = 0$, $y_2 = 0$, and leads to similar results in the case of other non-analytic surfaces. The theory of integral operators indicates that the above theorems represent only special cases of relations which exist for functions whose coefficients satisfy relation of the type indicated in (8. 1). In particular, from the integral representation of functions satisfying a differential equation of order $2n$, $n > 1$, follow similar theorems for solutions of linear differential equations of higher order. The author hopes in the future to return to the investigations of these questions in greater detail.

DIVISION OF ENGINEERING SCIENCES,
HARVARD UNIVERSITY.

[26] These considerations may also be regarded as a generalization of investigations of solutions of ordinary differential equations with singular coefficients (Fuchs theory) to the case of two variables, in which case a partial differential equation arises instead of an ordinary one.

BIBLIOGRAPHY

[1] Ludwig Baumgarten, "Beiträge zur Theorie der ganzen Funktionen von zwei komplexen Veränderlichen," *Monatshefte für Mathematik und Physik*, vol. 25 1914), pp. 3-70.

[2] Stefan Bergman, "Sur les fonctions entières et meromorphes de deux variables complexes," *Compositio Mathematica*, vol. 3 (1936), pp. 136-173.

[3] ———, "Zur Theorie der Funktionen, die eine lineare partielle Differentialgleichung befriedigen, I," *Recueil Mathematique*, new series, vol. 2 (1937), pp. 1169-1198.

[4] ———, "A formula for the stream function of certain flows," *Proceedings of the National Academy of Sciences*, vol. 29 (1943), pp. 276-281.

[5] ———, "Linear operators in the theory of partial differential equations," *Transactions of the American Mathematical Society*, vol. 53 (1943), pp. 130-155.

[6] ———, "Certain classes of analytic functions of two real variables and their properties," *Transactions of the American Mathematical Society*, vol. 57 (1945), pp. 299-331.

[7] ———, "On two-dimensional flows of compressible fluids," *N. A. C. A. Technical Note 972* (1945).

[8] ———, "Two-dimensional subsonic flows of a compressible fluid and their singularities," *Transactions of the American Mathematical Society*, vol. 62 (1947), pp. 452-498.

[9] ———, "Two-dimensional transonic flow patterns," *American Journal of Mathematics*, vol. 70 (1948), pp. 856-891.

[10] ———, "An initial value problem for equations of mixed type," *Bulletin of the American Mathematical Society*, vol. 56 (1949), pp. 165-174.

[11] ———, "On tables for the determination of transonic flow patterns," *Hans Reissner Anniversary Volume*, Brooklyn Polytechnic Institute (1949), pp. 13-36.

[12] L. Bers and A. Gelbart, "On a class of differential equations in mechanics of continua," *Quarterly of Applied Mathematics*, vol. 1 (1943), pp. 168-188.

[1]3 ———, "On a class of functions defined by partial differential equations," *Transactions of the American Mathematical Society*, vol. 56 (1944), pp. 67-93.

[14] Herman Chernoff, "Complex solutions of partial differential equations," *American Journal of Mathematics*, vol. 68 (1946), pp. 455-478.

[15] E. Holmgren, "Sur un probleme aux limites pour l'equation $y^m \partial^2 z/\partial x^2 + \partial^2 z/\partial y^2 = 0$," *Arkiv fur Mathematik, Astronomi och Fysik*, vol. 19, B (1927), pp. 1-3.

[16] Jules Sire, "Sur les fonctions entières de deux variables d'ordre apparent total fini," *Rendiconti del Circolo Matematico di Palermo*, vol. 31 (1910), pp. 1-91.

[17] ———, "Sur les fonctions entières de deux variables d'ordre apparent total fini et a croissance regulière par rapport a l'une des variables," *Journal de Mathématiques pures et appliqués*, vol. 9 (1913), pp. 1-37.

[18] F. Tricomi, "Sulle equazioni lineari alle derivato parziali di 2° ordina di tipo misto," *Memorie della Reale Accademia Nazionale dei Lincei*, vol. 14 (1923), pp. 133-247.

ON THE APPLICABILITY OF LINEAR DIFFERENTIAL OPERATORS OF INFINITE ORDER TO FUNCTIONS OF CLASS $L^2(B)$.*

By PHILIP DAVIS.

1. Introduction. In the plane of the complex variable $z = x + iy$, we shall designate by $L^2(B)$ the class of functions $f(z)$ which are regular and single valued in a fixed domain B and such that

$$\iint_B |f(z)|^2 dxdy < \infty. \tag{1}$$

By introducing the integral

$$(f, g) = \iint_B f\bar{g}\, dxdy \tag{2}$$

as an inner product, and $\| f \|^2 = (f, f)$ as a norm, the class $L^2(B)$ becomes a Hilbert space. This space is intimately related to problems in conformal mapping, and has been studied extensively. Reference is made to Bergman [2] where a systematic theory will be found.

In the present paper we study the class of bounded linear functionals defined on $L^2(B)$ considered as differential operators of infinite order. We show that such an operator is generated by an entire function of exponential type, and is a member of a class of functions, designated by $E^2(B)$ which are transforms of the functions of $L^2(B)$ arising from a Laplace kernel. The properties of $E^2(B)$ are investigated, and there is obtained a generalization of a theorem due to Muggli [7] on the applicability of linear differential operators where the case of functions analytic in a circle has been considered. In later sections, the inversion of this transform is discussed, and these considerations are applied to the question of the completeness of sets $\{e^{\lambda_n z}\}$ for $L^2(B)$, and to the coefficient problem for functions of exponential type with a preassigned indicator. This extends a familiar theorem of Lindelöf.

We shall find it convenient to recall a few basic facts about the space $L^2(B)$. To every B, there corresponds a complete orthonormal system of functions $\{\phi_n(z)\}$, and a "strong" Riesz-Fischer Theorem holds. Moreover, there is a function of two complex variables $K_B(z, \bar{t})$, known as the Bergman

* Received January 22, 1951; revised July 12, 1951.

kernel of the domain, which has proved to be of prime importance in the theory. The kernel belongs to $L^2(B)$ in both z and $\bar{t}$, and may be defined by

$$K_B(z,\bar{t}) = \sum_{n=0}^{\infty} \phi_n(z)\,(\phi_n(t))^- \tag{3}$$

but is independent of the particular choice of complete orthonormal system used in (3). Here $(\phi_n(t))^-$ designates the complex conjugate of $\phi_n(t)$.

In the present paper we shall restrict our domains B as follows. It will be required that B be schlicht, bounded, and simply connected. Moreover, if the boundary of B is b, it will be required that all the points of the plane which do not lie in $B + b$ form a single region whose boundary is precisely b. Thus a slit circle would be ruled out. It has been shown by Farrell [5] that for such domains, the set of powers is complete. If, therefore, B is such a domain, there exists a unique complete set of orthonormal polynomials

$$p_n(z) = \sum_{k=0}^{n} a_{nk} z^k; \qquad \mathrm{Re}(a_{nn}) \geqq 0,\ a_{nn} \neq 0, \tag{4}$$

where the coefficients a_{nk} have been determined by the familiar Gram-Schmidt orthonormalization process. Although a number of our theorems go through for more general domains, we shall find it convenient to restrict B as above so that a complete orthonormal set of polynomials is available.

In addition to the class $L^2(B)$, we shall be dealing with linear differential operators of infinite order

$$L(d) = \sum_{n=0}^{\infty} \beta_n d^n, \tag{5}$$

where the symbolic d^n is thought of as having the force of an n-th derivative evaluated at the origin: $L(d)f = \sum_{n=0}^{\infty} \beta_n f^{(n)}(0)$. It is convenient to replace the d in (5) by the complex variable z and to think of the operator $L(d)$ as being generated by the formal power series [1] $L(z) = \sum_{n=0}^{\infty} \beta_n z^n$. Our initial aim is to distinguish a class of operators (5) which will be applicable to functions of $L^2(B)$, and to characterize their generating power series.

2. On the applicability of linear differential operators. At the outset, we make the following definition. Suppose that $L(d)$ is given by (5), and that $f(z) \in L^2(B)$ has the Fourier expansion

[1] On the utility of this device for the discussion of certain problems arising in interpolatory function theory, see, e. g., Boas [3], Buck [4].

$$f(z) = \sum_{n=0}^{\infty} \alpha_n p_n(z), \qquad \alpha_n = (f, p_n). \tag{6}$$

We shall say that $L(d)$ is *applicable* (B) to f if and only if the series

$$\sum_{n=0}^{\infty} \alpha_n \sum_{k=0}^{n} \beta_k p_n^{(k)}(0) = \sum_{n=0}^{\infty} \alpha_n \sum_{k=0}^{n} a_{nk}\beta_k k! \tag{7}$$

converges. If the sum in (7) is r, then we shall write $L(d)_B f = r$.

Let us observe that if $p(z)$ is a polynomial of degree N, then $p(z) = \sum_{n=0}^{N} \alpha_n p_n(z)$, so that

$$\begin{aligned} L(d)_B p &= \sum_{n=0}^{N} \alpha_n \sum_{k=0}^{n} a_{nk}\beta_k k! = \sum_{k=0}^{N} \beta_k \sum_{k=n}^{N} \alpha_n a_{nk} k! \\ &= \sum_{k=0}^{\infty} \beta_k p^{(k)}(0) = L(d)p. \end{aligned} \tag{8}$$

Moreover, if $L(z)$ has a positive radius of convergence, and if $p(z) = \sum_{n=0}^{N} c_n z^n$, then we shall also have

$$L(d)_B p = \sum_{k=0}^{\infty} \beta_k p^{(k)}(0) = \sum_{k=0}^{\infty} \beta_k c_k k! = \sum_{k=0}^{N} c_k L^{(k)}(0) = p(d)L(z). \tag{9}$$

By virtue of this observation, we may say that $L(d)$ is applicable (B) to $f(z)$ if and only if $\sum_{n=0}^{\infty} \alpha_n L(d)p_n(z)$ converges.

THEOREM 1. *$L(d)$ is applicable (B) to all functions of $L^2(B)$ if and only if*

$$\sum_{n=0}^{\infty} \left| \sum_{k=0}^{n} a_{nk}\beta_k k! \right|^2 < \infty. \tag{10}$$

In this case, the operator is bounded.

Proof. By our definition, the series $\sum_{n=0}^{\infty} \alpha_n \sum_{k=0}^{n} a_{nk}\beta_k k!$ is required to converge for all sequences $\{\alpha_n\}$ with $\sum_{n=0}^{\infty} |\alpha_n|^2 < \infty$. A familiar theorem of Landau [6] tells us that (10) is a necessary and sufficient condition for this. If we set

$$K^2 = \sum_{n=0}^{\infty} \left| \sum_{k=0}^{n} a_{nk}\beta_k k! \right|^2 \tag{11}$$

then, by the Schwarz inequality, $|L(d)_B f|^2 \leqq K^2 \sum_{n=0}^{\infty} |\alpha_n|^2 = K^2 \|f\|^2$, establishing the boundedness of the operator.

For the circle $C_r: |z| < r$, the orthonormal polynomials are given explicitly by

$$p_n(z) = \pi^{-\frac{1}{2}}(n+1)^{\frac{1}{2}} z^n / r^{n+1} \qquad (n = 0, 1, \cdots), \tag{12}$$

and as a result of their simple structure, we see from (7) that if $L(d)$ is applicable to $f(z) \in L^2(C_r)$, then $L(d)_{C_r} f = L(d)f$. Thus for the case of the circle, applicability as defined by (7) reduces to ordinary applicability (cf. also (49)). For the circle, the criterion (10) becomes

$$\sum_{n=0}^{\infty} |\beta_n|^2 n!(n+1)!r^{-2n} < \infty. \tag{13}$$

In his connection, cf. Muggli [7].

For a fixed domain B, we shall designate the class of power series $L(z)$ whose coefficients $\{\beta_n\}$ satisfy (10) by $E^2(B)$. It will appear shortly that the power series in this class have more than formal meaning; indeed, they are entire functions of exponential type. The exact scope of the linear differential operators of class $E^2(B)$ is indicated by the following converse of Theorem 1.

THEOREM 2. *Any bounded linear functional defined on $L^2(B)$ may be represented as a linear differential operator of infinite order and of class $E^2(B)$.*

Proof. From the general theory (see, e. g., Stone [9], p. 62) it is known that if L is such a functional, there will exist an $h(z) \in L^2(B)$ such that

$$Lf = (f, h) \qquad \text{for all } f \in L^2(B). \tag{14}$$

Suppose that $h(z)$ has the Fourier expansion $h(z) = \sum_{n=0}^{\infty} \alpha_n p_n(z)$. Set up the infinite system of linear equations in the unknowns β_k $(k = 0, 1, \cdots)$,

$$\sum_{k=0}^{n} a_{nk}\beta_k k! = \alpha_n \qquad (n = 0, 1, \cdots). \tag{15}$$

Inasmuch as the system (15) is triangular, and $a_{nn} \neq 0$, it will be possible to arrive at a solution in a step-by-step fashion. Since $\sum_{n=0}^{\infty} |\alpha_n|^2 < \infty$, the power series $L(z) = \sum_{k=0}^{\infty} \beta_k z^k$ has the property (10). Thus, $L(z) \in E^2(B)$, and $L(d)$ is applicable (B) to all $f(z) \in L^2(B)$. If now $f(z) = \sum_{n=0}^{\infty} \gamma_n p_n(z)$, then $L(d)_B f(z) = \sum_{n=0}^{\infty} \gamma_n \sum_{k=0}^{n} a_{nk}\beta_k k!$. But

$$(f, h) = \iint_B \sum_{n=0}^{\infty} \gamma_n p_n(z) \Big(\sum_{m=0}^{\infty} \alpha_m p_m(z) \Big)^- dxdy$$

$$= \sum_{n=0}^{\infty} \gamma_n \alpha_n = \sum_{n=0}^{\infty} \gamma_n \sum_{k=0}^{n} a_{nk} \beta_k k! = L(d)_B f.$$

We next give an explicit construction for the function $h(z)$ associated with the functional $L(z)$ through the relation (14). It is obtained by operating on one variable of the kernel $K_B(z, \bar{t})$ by $L(d)$. In this connection, see also Aronszajn [1]. We shall merely state the result.

THEOREM 3. *Let $L(z) \in E^2(B)$. Then the function*

$$h(t) = \sum_{n=0}^{\infty} (L(d) p_n(z))^- p_n(t) = (L(d_z)_B K_B(z, \bar{t}))^- \tag{16}$$

belongs to $L^2(B)$, and is such that

$$L(d)_B f = (f, h) \tag{17}$$

for all $f \in L^2(B)$. The subscript in $L(d_z)$ indicates that the operation is to be performed on the variable z.

The quantity $(h, h)^{\frac{1}{2}}$ is the norm of the operator, and we have explicitly

$$\| L(d) \|^2 = \sum_{n=0}^{\infty} | L(d) p_n(z) |^2 = \sum_{n=0}^{\infty} | \sum_{k=0}^{n} a_{nk} \beta_k k! |^2. \tag{18}$$

The norm of $L(d)$ depends, of course, upon the domain B, and when we wish to emphasize the role of B we shall write $\| L(d) \|_B$.

3. The class $E^2(B)$. In this section we characterize more closely the class of power series $E^2(B)$ whose Maclaurin coefficients $\{\beta_n\}$ are subject to (10). First we shall find a representation for functions of this class, showing that it is identical with the class of transforms of $f \in L^2(B)$ arising from a Laplace kernel.

THEOREM 4. *$g(z) \in E^2(B)$ if and only if*

$$g(z) = \iint_B e^{zw} (f(w))^- dudv; \qquad w = u + iv, \tag{19}$$

with $f(z) \in L^2(B)$.

Proof. Let $f(z) \in L^2(B)$ and have Fourier coefficients $\alpha_n = (f, p_n)$.

Then we have

$$g(z) = \int\int_B e^{zw}(f(w))^- du dv = \int\int_B \sum_{n=0}^{\infty} z^n w^n (f(w))^-/n!\, du dv$$
$$= \sum_{n=0}^{\infty} \beta_n z^n,$$

where $\beta_n = (n!)^{-1}(w^n, f(w))$. Therefore,

$$\sum_{k=0}^{n} a_{nk}\beta_k k! = \sum_{k=0}^{n} a_{nk}(w^n, f(w)) = (p_n, f) = \alpha_{\bar{n}},$$

so that condition (10) holds. Conversely, given a $g(z) = \sum_{n=0}^{\infty} \beta_n z^n$ with (10) holding, there exists by the Riesz-Fischer Theorem an $f(z) \in L^2(B)$ such that the quantities $(\sum_{k=0}^{n} a_{nk}\beta_k k!)^-$ are its Fourier coefficients:

$$\sum_{k=0}^{n} a_{nk}\beta_k k! = (p_n, f) \qquad (n = 0, 1, \cdots). \tag{20}$$

We now claim that $g(z) = \sum_{k=0}^{\infty} \beta_k z^k = \int\int_B e^{zw}(f(w))^- du dv$. To show this, it suffices, as above, to prove that

$$\beta_k k! = (w^k, f(w)) \qquad (k = 0, 1, \cdots). \tag{21}$$

From (20), we have

$$\sum_{k=0}^{n} a_{nk}\beta_k k! = (\sum_{k=0}^{n} a_{nk}w^k, f(w)) = \sum_{k=0}^{n} a_{nk}(w^k, f(w)), \tag{22}$$

$(n = 0, 1, \cdots)$. This triangular system may now be solved step-by-step, and yields (21) as its solution.

It will be convenient to abbreviate the transform (19) as $g = \mathfrak{L}_B(f)$, and when this notation is employed, it will be understood that $f \in L^2(B)$. We shall now characterize the growth of the transforms $\mathfrak{L}_B(f)$, but before doing so, it will be useful to recall a number of definitions and theorems which are of importance in the theory of entire functions of exponential type. For this material, reference is made to Pólya [8]. By the Phragmén-Lindelöf indicator of an entire function of exponential type $f(z)$ in the direction θ is meant the quantity

$$h(\theta) = \limsup_{r\to\infty} r^{-1} \log |f(re^{i\theta})| \tag{23}$$

If we have $h(\theta) \leqq \sigma$ for $0 \leqq \theta \leqq 2\pi$, we shall speak of an entire function

of exponential type at most σ, a class designated by $(1, \sigma)$. If $f(z) = \sum_{n=0}^{\infty} a_n z^n$ is an entire function of exponential type, then the associated series

$$\mathfrak{B}(f) = \sum_{n=0}^{\infty} a_n n!/z^{n+1}, \tag{24}$$

is known as the Borel transform of f. An entire function $f(z)$ is of class $(1, \sigma)$ if and only if the radius of convergence of $\mathfrak{B}(f)$ is at most σ. The closed convex hull of the set of singularities of $\mathfrak{B}(f)$, known as the conjugate indicator diagram of f, will be designated by $D(f)$. If $f \in (1, \sigma)$, then it possesses a representation of the form

$$f(z) = (2\pi i)^{-1} \int_C e^{zw} \mathfrak{B}f(w)\, dw, \tag{25}$$

where C is any simple closed curve containing $D(f)$ in its interior. If B is a bounded convex domain, its supporting function $k(\theta)$ is defined by

$$k(\theta) = \max_{(x,y) \in \bar{B}} (x \cos\theta + y \sin\theta). \tag{26}$$

The indicator of an entire function of exponential type and the supporting function of its conjugate indicator diagram are related intimately by the formula

$$h(\theta) = k(-\theta) \qquad 0 \leqq \theta \leqq 2\pi. \tag{27}$$

THEOREM 5. *Let $\hat{B}$ designate the closed convex hull of a given domain B. Every function $g(z)$ of the class $E^2(B)$ is an entire function of exponential type. If $h(\theta)$ is its indicator, and if $k(\theta)$ is the supporting function of $\hat{B}$, then*

$$h(\theta) \leqq k(-\theta), \qquad 0 \leqq \theta \leqq 2\pi. \tag{28}$$

Furthermore, we have $D(g) \subseteq B$.

Proof. With $g(z) \in E^2(B)$, we have, by Theorem 4, $g = \mathfrak{L}_B(f)$ for some $f \in L^2(B)$. Applying the Schwarz inequality, we obtain

$$\begin{aligned} |g(re^{i\theta})|^2 &\leqq \|f\|^2 \iint_B |\exp(re^{i\theta}w)|^2\, du dv \\ &= \|f\|^2 \iint_B \exp\{2r(u\cos\theta - v\sin\theta)\}\, du dv. \end{aligned} \tag{29}$$

Now for $(u, v) \in \hat{B}$, and hence for $(u, v) \in B$, we have by (26)

$$u\cos\theta - v\sin\theta \leqq k(-\theta), \qquad 0 \leqq \theta \leqq 2\pi,$$

so that (29) becomes $|g(re^{i\theta})|^2 \leqq \|f\|^2 \exp\{2rk(-\theta)\}$ area (B). Therefore,

$h(\theta) = \limsup_{r\to\infty} r^{-1} \log |g(re^{i\theta})| \leqq k(-\theta)$. Finally, by (27), $h(-\theta)$ is the supporting function of $D(g)$. By virtue of the last inequality, we have $h(-\theta) \leqq k(\theta)$, and this implies that $D(g) \subseteq B$.

While $D(g) \subseteq B$, the former need not be interior to the latter. We shall give an example which shows that in certain cases the boundary of $\hat{B}$ may contain a boundary point of $D(g)$. Take B as the unit circle. For this domain $\hat{B} = \hat{B}$, and $(z^m, z^n) = \pi\delta_{mn}/(n+1)$ $(m, n = 0, 1, \cdots)$. Consider the function $\phi(z) = \pi^{-1} \sum_{n=0}^{\infty} z^n/(n+1)^{\frac{1}{2}}$. ϕ is regular in the unit circle, and $\| \phi \|^2 = \sum (n+1)^{-2} < \infty$ (cf. (12)), so that $\phi \in L^2(B)$. Now $\psi(z) = \mathfrak{L}_B(\phi) = \sum_{n=0}^{\infty} z^n/n!(n+1)^{3/2}$, so that $\mathfrak{B}(\psi) = \sum_{n=0}^{\infty} (n+1)^{-3/2} z^{-n-1}$. This last series converges at points exterior to B, and $\mathfrak{B}(\psi)$ possesses a singularity on the circumference of B.

A basis for the space $E^2(B)$ is provided by the functions

$$q_n(z) = \iint_B e^{zw} (p_n(w))^- \, du dv \qquad (n = 0, 1, \cdots). \tag{30}$$

By Theorem 5, the q's are entire of exponential type. Since also

$$e^{zw} = \Big(\sum_{k=0}^{n-1} + \sum_{k=n}^{\infty}\Big) z^k w^k / k!,$$

and each $p_n(z)$ is orthogonal to all polynomials of lower degree, we have

$$q_n(z) = b_{nn} z^n/n! + b_{n+1n} z^{n+1}/(n+1)! + \cdots \qquad (n = 0, 1, \cdots). \tag{31}$$

It is easily shown that all functions of the class $E^2(B)$ may be expanded in a series of q's, convergent uniformly in every bounded domain.

4. Inversion formulas for the $\mathfrak{L}_B$ transform. In the present section, we derive a number of formulas for inverting the $\mathfrak{L}_B$ transform (19).

THEOREM 6. *Let $g = \mathfrak{L}_B(f)$. Then*

$$f(z) = \sum_{n=0}^{\infty} (g(d) p_n(z))^- p_n(z), \tag{32}$$

$$f(z) = \sum_{n=0}^{\infty} (g(d_w)\, {}_B K_B(w, \bar{z}))^-, \tag{33}$$

$$f(z) = \sum_{n=0}^{\infty} p_n(z) \{(2\pi i)^{-1} \int_C g(z) \mathfrak{B}(p_n(z))\, dz\}^-, \tag{34}$$

where C is any simple closed curve containing the origin in its interior. If it is assumed, in addition, that $D(g)$ lies in the interior of B, then we have

$$(35) \qquad (f(z))^- = (2\pi i)^{-1} \int_{C_1} K_B(w, \bar{z}) \mathfrak{B}(g(w))\, dw,$$

where C_1 is any simple closed curve lying in the interior of B and containing $D(g)$ in its interior.

Proof. With $p_n(z) = \sum_{k=0}^{n} a_{nk} z^k$, we have by (9)

$$(36) \qquad g(d)p_n(z) = \sum_{k=0}^{n} a_{nk} g^{(k)}(0) = \sum_{k=0}^{n} a_{nk}(w^k, f) = (p_n, f).$$

Hence the Fourier coefficients of f are $(g(d)p_n(z))^-$, and (32) follows. Equation (33) follows from (32) by Theorem 3. To arrive at (34), we notice that by Cauchy's theorem and (24),

$$(37) \qquad (2\pi i)^{-1} \int_C g(z) \mathfrak{B}(p_n(z))\, dz = \sum_{k=0}^{n} a_{nk} g^{(k)}(0) = (p_n, f).$$

To establish (35), we proceed as follows. Let C designate a circle centered at the origin and of radius sufficiently large so that $D(g)$ is contained in its interior. If $g(z) = \sum_{n=0}^{\infty} \gamma_n z^n$, then the function $\mathfrak{B}(g(z)) = \sum_{n=0}^{\infty} \gamma_n n!/z^{n+1}$ is regular on C_1, while its Laurent series converges absolutely and uniformly on C. Thus we have

$$\begin{aligned}(2\pi i)^{-1} \int_{C_1} p_n(z) \mathfrak{B}(g(z))\, dz &= (2\pi i)^{-1} \int_C p_n(z) \mathfrak{B}(g(z))\, dz \\ &= (2\pi i)^{-1} \int_C p_n(z) \sum_{k=0}^{\infty} \gamma_k k!/z^{k+1} dz = \sum_{k=0}^{n} a_{nk} \gamma_k k! \\ &= \sum_{k=0}^{n} a_{nk} g^{(k)}(0) = (2\pi i)^{-1} \int_C g(z) \mathfrak{B}(p_n(z))\, dz.\end{aligned}$$

The last equality follows by (37). By (34), therefore,

$$(f(z))^- = \sum_{n=0}^{\infty} (p_n(z))^- (2\pi i)^{-1} \int_{C_1} p_n(z) \mathfrak{B}(g(z))\, dz.$$

Since $K_B(w, \bar{z})$ is given by (3), and for fixed z converges uniformly for w in any closed subset of B, we arrive at (35) easily.

COROLLARY. *Let $g = \mathfrak{L}_B(f)$. Then for all $h(z) \in L^2(B)$,*

$$(38) \qquad g(d)_B h(z) = (h, f).$$

Proof. This follows directly from Theorems 3 and 6.

COROLLARY. *If* $g(z) \equiv 0$, *then* $f(z) \equiv 0$.

Two functions whose $\mathfrak{Q}_B$ transforms are identical are therefore themselves identical. It should be observed that this result has been derived on the assumption that B is a domain for which the powers of z form a complete system. It will be shown in a later paragraph that there exist bounded simply-connected domains for which the $\mathfrak{Q}_B$ transformation is not one-to-one.

We next record a number of identities relating the orthonormal polynomials and their $\mathfrak{Q}_B$ transforms $q_n(z)$.

COROLLARY. *The sets* $\{p_n(z)\}$ *and* $\{q_n(z)\}$ *are biorthonormal in the sense that*

$$q_m(d)p_n(z) = p_n(d)q_m(z) = \delta_{mn} \qquad (m, n = 0, 1, \cdots). \tag{39}$$

This results from the first corollary above and the orthonormality of the p_n's. The biorthonormality relations (39) may also be expressed in the following two ways. Let A and B be the ∞ by ∞ matrices of coefficients

$$A = (a_{nk}), \qquad B = (b_{nk}), \tag{40}$$

with a_{nk} given by (4) and b_{nk} by (31). Then

$$AB = BA = I. \tag{41}$$

Alternately, the following identity holds for w arbitrary and uniformly for z lying in a closed subdomain of B.

$$e^{zw} = \sum_{n=0}^{\infty} q_n(w)p_n(z). \tag{42}$$

In section 3 we proved that if $g(z) \varepsilon E^2(B)$, then $D(g) \subseteq \hat{B}$. In addition, we gave an example showing that the boundaries of $D(g)$ and of $\hat{B}$ may have points in common. In the reverse direction, we shall show in Theorem 7 that if $D(g)$ is contained in the interior of B, then $g(z)$ is surely a member of $E^2(B)$. But if we know only that $D(g) \subseteq \hat{B}$, it may not follow that $g \varepsilon E^2(B)$.

As an example of this behavior, let B be selected as the unit circle. For $g(z)$, choose the entire function $g(z) = (e^z - 1)/z$. It is easily verified that $\mathfrak{B}(g) = \log(1 - z^{-1})$. Therefore $D(g)$ has a point which lies on the boundary of B, viz. $z = 1$. Again, we have $p_n(z) = (\pi)^{-\frac{1}{2}}(n+1)^{\frac{1}{2}}z^n$, and assuming that $g = \mathfrak{Q}_B(f)$, where $f = \sum_{n=0}^{\infty} \alpha_n p_n$, we find that $\alpha_n = (\pi(n+1))^{-\frac{1}{2}}$. Thus $\| f \| = \infty$, and g is the transform of no $f \varepsilon L^2(B)$.

THEOREM 7. *Let $g(z)$ be an entire function of exponential type and such that $D(g) \subset B$. Then $g(z) \in E^2(B)$.*

Proof. Let C_1 be a simple closed rectifiable contour lying in B and containing $D(g)$ in its interior. Construct the function

$$f(z) = \sum_{n=0}^{\infty} k_n p_n(z), \tag{43}$$

where

$$k_n = (2\pi i)^{-1} \int_{C_1} p_n(z) \mathfrak{B}(g(z))\, dz \qquad (n = 0, 1, \cdots). \tag{44}$$

Applying the Schwarz inequality to (44), we obtain

$$(2\pi)^2 |k_n|^2 \leqq l(C_1) \max_{z \in C_1} |\mathfrak{B}(g(z))|^2 \int_{C_1} |p_n(z)|^2\, ds,$$

so that

$$(2\pi)^2 \sum_{n=0}^{\infty} |k_n|^2 \leqq l(C_1) \max_{z \in C_1} |\mathfrak{B}(g(z))|^2 \int_{C_1} K_B(z, \bar{z})\, ds < \infty,$$

inasmuch as C_1 is interior to B. Therefore $f \in L^2(B)$, and we have

$$\begin{aligned} g(z) &= (2\pi i)^{-1} \int_{C_1} e^{zw} \mathfrak{B}(g(w))\, dw = (2\pi i)^{-1} \int_{C_1} \sum_{n=0}^{\infty} q_n(z) p_n(w) \mathfrak{B}(g(w))\, dw \\ &= \sum_{n=0}^{\infty} q_n(z) k_n = \mathfrak{L}_B(f(z)) \end{aligned}$$

by (25), (42), (43), and (30) successively.

COROLLARY. *Let B_1 and B_2 be two domains such that $\hat{B}_1 \subset B_2$. Then $E^2(B_1) \subset E^2(B_2)$. In particular, the conclusion holds if $B_1 \subset B_2$ and at least one of the two domains is convex.*

Proof. Let $g(z) \in E^2(B_1)$. Then $D(g) \subseteq \hat{B}_1 \subset B_2$ by Theorem 5 and our hypothesis. Therefore, by the present theorem, $g(z) \in E^2(B_2)$.

Under the hypotheses of Theorem 7, an alternate expression for $g(d)_B f(z)$, $f \in L^2(B)$, may be found. From (38) and (35) we obtain

$$g(d)_B f(z) = \iint_B f(z) (2\pi i)^{-1} \int_{C_1} K_B(w, \bar{z}) \mathfrak{B}(g(w))\, dw\, dx dy, \tag{45}$$

where C_1 is a simple closed curve lying in the interior of B and containing $D(g)$ in its interior. Thus we have

$$\begin{aligned} g(d)_B f(z) &= (2\pi i)^{-1} \int_{C_1} \mathfrak{B}(g(w))\, dw \iint_B f(z) K_B(w, \bar{z})\, dx dy \\ &= (2\pi i)^{-1} \int_{C_1} f(w) \mathfrak{B}(g(w))\, dw. \end{aligned} \tag{46}$$

If it is further assumed that B is a star domain with respect to the origin, then by the well-known theory of Mittag-Leffler summability of Taylor series we have

$$f(w) = \lim_{\delta \to \infty} \sum_{n=0}^{\infty} f^{(n)}(0) w^n / n! \Gamma(1 + (n-1)\delta), \tag{47}$$

uniformly on every closed subset of B. With (46) this yields

$$\begin{aligned} g(d)_B f &= \lim_{\delta \to 0} \sum_{n=0}^{\infty} (2\pi i)^{-1} \int_{C_1} w^n \mathfrak{B}(g(w))\, dw\, f^{(n)}(0)/n!\Gamma(1+(n-1)\delta) \\ &= \lim_{\delta \to 0} \sum_{n=0}^{\infty} a_n f^{(n)}(0)/\Gamma(1+(n-1)\delta), \end{aligned} \tag{48}$$

where we have assumed that $g(z) = \sum_{n=0}^{\infty} a_n z^n$. This may be written

$$g(d)_B f(z) = (M) \sum_{n=0}^{\infty} a_n f^{(n)}(0), \tag{49}$$

the symbol (M) indicating a process of Mittag-Leffler summation. If it is known from independent considerations that $\sum_{n=0}^{\infty} a_n f^{(n)}(0)$ converges, then since (M) is a regular method of summation, we have

$$g(d)_B f(z) = \sum_{n=0}^{\infty} a_n f^{(n)}(0) = g(d) f(z). \tag{50}$$

5. Completeness of the set $\{e^{\lambda_n z}\}$. In the present section, we employ the inversion formula just derived plus the well-known device of Szász [10] to obtain conditions on the distribution of the complex values λ_n under which the set of functions $\{e^{\lambda_n z}\}$ $(n = 0, 1, \cdots)$ will be complete for $L^2(B)$. By this device, the completeness is referred to the possible distribution of zeros of the entire functions of class $E^2(B)$.

THEOREM 8. *Let $\lambda_0, \lambda_1, \cdots$ be a sequence of distinct complex numbers with $\lambda_n \to \infty$. Let $n(x)$ denote the number of λ's for which $|\lambda_n| \leqq x$. Let $c(\hat{B})$ designate the circumference of the convex hull of B. If*

$$\limsup_{r \to \infty} r^{-1} \int_1^r (n(x)/x)\, dx > c(\hat{B})/2\pi, \tag{51}$$

then the set $\{e^{\lambda_n z}\}$ will be complete in $L^2(B)$.

Proof. Suppose that the sequence of λ's satisfies (51) but that the set is not complete. Then there will exist an $f \in L^2(B)$ not identically zero, such that $(e^{\lambda_n z}, f(z)) = 0$ $(n = 0, 1, \cdots)$. Consider now the function $g = \mathfrak{L}_B(f)$.

The entire function $g(z)$ is not identically zero by the second corollary to Theorem 6. The points $z = \lambda_n$ form a subset of its zeros. If, therefore, $n^*(x)$ denotes the number of zeros of g whose moduli do not exceed x, we shall have $n(x) \leqq n^*(x)$. By Jensen's Theorem (see, e. g., Titchmarsh (11)) we have

$$(52) \quad \int_1^r (n(x)/x)\,dx \leqq \int_1^r (n^*(x)/x)\,dx < (2\pi)^{-1} \int_0^{2\pi} \log |\, g(re^{i\theta})|\, d\theta + c,$$

where c is a fixed constant which depends solely upon g. Now from Theorem 5 we have $\log |\, g(re^{i\theta})| \leqq rk(-\theta) + \log \| f \| + \frac{1}{2} \log \text{Area } (B)$, so that

$$\limsup_{r\to\infty} r^{-1} \int_1^r (n(x)/x)\,dx \leqq (2\pi)^{-1} \int_0^{2\pi} k(-\theta)\,d\theta = c(\hat{B})/2\pi$$

(on this last equality see [8]). This contradicts (51).

It is possible to show that for convex domains possessing a center of symmetry, the constant $c(\hat{B})/2\pi$ is the best possible. This fact falls out as a corollary to a more general theorem which we shall now prove.

THEOREM 9. *Let the domain B contain in its interior a convex polygon P which posesses a center of symmetry. Then there exist systems $\{e^{\lambda_n z}\}$ with*

$$(53) \qquad \lim_{r\to\infty} r^{-1} \int_1^r (n(x)/x)\,dx = c(P)/2\pi$$

which are not complete in $L^2(B)$.

Proof. We assume that $z = 0$ is the center of symmetry of P. Select a vertex of P, denote it by z_1, and, proceeding counterclockwise, let the remaining vertices be denoted by $z_2, z_3, \cdots, z_n, -z_1, -z_2, \cdots, -z_n$. Set

$$(54) \qquad \begin{aligned} &w_1 = \tfrac{1}{2}(z_2 - z_1), \quad w_2 = \tfrac{1}{2}(z_3 - z_2), \cdots, w_{n-1} = \tfrac{1}{2}(z_n - z_{n-1}), \\ &w_n = \tfrac{1}{2}(-z_1 - z_n). \end{aligned}$$

Then the circumference of P is $c(P) = 4(|\, w_1 \,| + |\, w_2 \,| + \cdots + |\, w_n \,|)$. Consider now the function

$$(55) \qquad \begin{aligned} \phi(z) &= \prod_{j=1}^{n} \sin(w_j z/i) = (-i/2)^n \prod_{j=1}^{n} (e^{w_j z} - e^{-w_j z}) \\ &= (-i/2)^n \sum \pm e^{\sum_{j=1}^{n} \pm w_j z} \end{aligned}$$

where the first sum is to be extended over the 2^n possible combinations of $\pm$ in the second sum. It follows from (24) that the singularities of $\mathfrak{B}(\phi(z))$

are located at the 2^n points $\sum_{j=1}^{n} \pm w_j$. Now it may be easily verified that the $2n$ vertices of P are located at points $\sum_{j=1}^{n} \pm w_j$ for some $2n$ selections of $\pm$, while the remaining $2^n - 2n$ selections give rise to points which are interior to P. Thus the conjugate indicator diagram of ϕ is precisely $P: D(\phi) = P$. Since P is interior to B, by Theorem 7, $\phi(z) = \mathfrak{L}_B(\psi)$ for some $\psi \varepsilon L^2(B)$. If $\lambda = m\pi i/w_j$ $(m = 0, 1, 2, \cdots; j = 1, 2, \cdots, n)$, then we have by (55), $(e^{\lambda z}, \psi) = 0$. Let λ_k denote these points arranged in order of non-decreasing absolute value. For this set it is clear that

$$(56) \qquad n(x) = (2x/\pi)(|w_1| + |w_2| + \cdots + |w_n|) + O(1), \quad x \to \infty.$$

Hence for the function ψ we have

$$\lim_{r\to\infty} (r^{-1}) \int_1^r (n(x)/x)\,dx = (2/\pi)(|w_1| + \cdots + |w_n|) = c(P)/2\pi.$$

COROLLARY. *If B is convex and possesses a center of symmetry, then there exist systems $\{e^{\lambda_n z}\}$ for which*

$$(57) \qquad \lim_{r\to\infty} (r^{-1}) \int_1^r (n(x)/x)\,dx = c(B)/2\pi - \epsilon,$$

and which are not complete in $L^2(B)$.

Proof. In this case we may find a polygon possessing the same center of symmetry, and lying in the interior of B, whose perimeter differs from that of B by arbitrarily small amounts.

Throughout the present work the assumption has been made that the domain B is such that the powers of z are complete for $L^2(B)$. Under this assumption, the $\mathfrak{L}_B$ transform has been shown to be one-to-one. However this transform may be defined for any bounded domain, and the order conditions of Theorem 5 will be preserved. But it is possible to show that in general the transform will not produce a one-to-one mapping of $L^2(B)$ on $E^2(B)$. To show this, suppose that the domain B, assumed to satisfy our usual conditions, is slit. That is, let a Jordan arc issuing from a boundary point of B and penetrating into its interior be added as a part of its boundary. Denote the resulting slit domain by B_s. The following theorem now holds.

THEOREM 10. *There exists a function $f(z) \varepsilon L^2(B_s)$ which is not identically zero, and is such that*

$$(58) \qquad \mathfrak{L}_{B_s}(f) \equiv 0.$$

Prooof. Assume that the conclusion is false and that the only function for which (58) holds is the null function. If a sequence of constants λ_n is selected so that condition (51) is satisfied, then the proof of Theorem 8 will go through word for word, and we will have proved that the set $\{e^{\lambda_n z}\}$ is complete for the class $L^2(B_s)$. Now this is impossible, for if we orthonormalize the set $\{e^{\lambda_n z}\}$ with respect to B_s, the resulting set of exponential polynomials will be complete and orthonormal for B itself, inasmuch as the slit does not affect the values of the area integrals which occur in the orthonormalization process. If, then, $f(z)\,\varepsilon\,L^2(B_s)$, it would have an expansion in orthogonal exponential polynomials converging in the interior of B, and we would have every function of $L^2(B_s)$ regular in B. This is a contradiction.

6. Application to a coefficient problem in the theory of entire functions of exponential type. Let $f(z)$ be an entire function of exponential type. As we have already mentioned, the indicator of f is essentially the supporting function of the convex domain $D(f): h_f(\theta) = k_{D(f)}(-\theta)$, $0 \leqq \theta \leqq 2\pi$. Conversely, it is known that for any given convex domain C there are entire functions f such that $D(f) = \bar{C}$. We now pose the following problem. Let $L(z) = \sum_{n=0}^{\infty} \beta_n z^n$. What are necessary and sufficient conditions on the sequence of coefficients $\{\beta_n\}$ in order that $L(z)$ have a preassigned indicator, or what is equivalent, in order that $D(L(z))$ be a given convex domain. Such a condition is conveniently given in terms of the norm of the operator $L(d)$.

THEOREM 11. *Let the domain B be convex and $k_B(\theta)$ denote its supporting function. In order that the indicator $h(\theta)$ of $L(z)$ satisfy*

$$h(\theta) = k_B(-\theta), \qquad 0 \leqq \theta \leqq 2\pi, \tag{59}$$

it is necessary and sufficient that

$$\| L(d) \|_{B^*} < \infty \qquad \text{for all } B^* \text{ with } \bar{B} \subset B^*, \tag{60}$$

and

$$\| L(d) \|_{B^*} = \infty \qquad \text{for all } B^* \text{ with } B \cap \mathrm{Ext} B^* \neq 0. \tag{61}$$

Proof. Let (60) and (61) hold. By (60), $L(z)\,\varepsilon\,E^2(B^*)$ for all B^* containing B in its interior. Select a sequence of convex domains B^*_n $(n = 1, 2, \cdots)$, such that $B^*_n \supseteq B^*_{n+1}$, and which converge to B. If $k_{B^*_n}(\theta)$ are the supporting functions of B^*_n then we have $\lim_{n\to\infty} k_{B^*_n}(\theta) = k_B(\theta)$, $0 \leqq \theta \leqq 2\pi$, so that since $h(\theta) \leqq k_{B^*_n}(-\theta)$, we have $h(\theta) \leqq k_B(-\theta)$ $(0 \leqq \theta \leqq 2\pi)$. If now $h(\theta) < k_B(-\theta)$ for some θ, then by the continuity

of $h(\theta)$, the inequality would hold over a range of values, and therefore B would have interior points which are exterior to $D(L(z))$. It would then be possible to find a domain B' containing $D(L(z))$ in its interior but such that B has points exterior to B'. Since $D(L(z))$ is interior to B', by Theorem 7, $L(z) \in E^2(B')$. This is impossible by (61). Therefore $h(\theta) = k_B(-\theta)$, $0 \leqq \theta \leqq 2\pi$.

Conversely, suppose that (59) holds. Then $D(L(z)) = \bar{B}$. Let B^* contain B in its interior; then by Theorem 7, $L(z) \in E^2(B^*)$, so that (60) must hold. Suppose that B contains an interior point exterior to a domain B^*, and that (61) does not hold. Then $L(z) \in E^2(B^*)$. From Theorem 5, $D(L(z))$ must lie in $\bar{B}^*$ and we have a contradiction.

If the orthonormal polynomials of a domain B^* are given by

$$p^*_n(z) = \sum_{k=0}^{n} a^*_{nk} z^k \qquad (n = 0, 1, \cdots),$$

then conditions (60) and (61) become

$$\sum_{n=0}^{\infty} \left| \sum_{k=0}^{n} a^*_{nk} \beta_k k! \right|^2 < \infty \text{ or } = \infty. \tag{62}$$

Since a^*_{nk} depend solely upon B^*, we see that (60), (61) are, in reality, conditions to be satisfied by the coefficients of $L(z)$.

As a special case, we may obtain the familiar Lindelöf relationship between the coefficients of the series development of an entire function of exponential type and the type. Let $g(z) = \sum_{n=0}^{\infty} \beta_n z^n$ be an entire function of exponential type σ. A necessary and sufficient condition for this is that $D(g)$ be contained in the circle $|z| \leqq \sigma$, and that at least one boundary point of $D(g)$ lie on $|z| = \sigma$. If C_r designates the circle $|z| = r$, then it is necessary and sufficient that

$$\| g(d) \|_{C_r} < \infty \quad \text{for} \quad r > \sigma, \tag{63}$$

$$\| g(d) \|_{C_r} = \infty \quad \text{for} \quad r < \sigma. \tag{64}$$

Referring to (13), $\| g(d) \|^2_{C_r} = \sum_{n=0}^{\infty} |\beta_n|^2 n!(n+1)!r^{-2n}$, while the radius of convergence of this series in r is

$$\sigma = \limsup_{m \to \infty} (m/e) |\beta_m|^{1/m}. \tag{65}$$

Theorem 11 can therefore be regarded as a generalization of Lindelöf's Theorem.

HARVARD UNIVERSITY.

BIBLIOGRAPHY.

[1] N. Aronszajn, "Theory of reproducing kernels," *Transactions of the American Mathematical Society*, vol. 68 (1950), pp. 337-404.

[2] S. Bergman, "The kernel function and conformal mapping," *Mathematical Surveys*, No. 5, New York, 1950.

[3] R. P. Boas, Jr., "Entire functions of exponential type," *Bulletin of the American Mathematical Society*, vol. 48 (1942), pp. 839-849.

[4] R. C. Buck, "Interpolation and uniqueness of entire functions," *Transactions of the American Mathematical Society*, vol. 64 (1948), pp. 283-298.

[5] O. J. Farrell, "On the approximation to an analytic function by polynomials," *Bulletin of the American Mathematical Society*, vol. 40 (1934), pp. 908-914.

[6] E. Landau, "Über einen Konvergenzsatz," *Nachrichten der Königlichen Gesellschaft der Wissenschaften zu Göttingen* (1907), pp. 25-27.

[7] H. Muggli, "Differentialgleichungen unendlichhoher Ordnung mit konstanten Koeffizienten," *Comentarii Mathematici Helvetici*, vol. 11 (1938), pp. 151-179.

[8] G. Pólya, "Untersuchungen über Lücken und Singularitaten von Potenzreihen," *Mathematische Zeitschrift*, vol. 29 (1929), pp. 549-640.

[9] M. H. Stone, "Linear transformations in Hilbert space," *American Mathematical Society Colloquium Publication*, vol. XV, New York, 1932.

[10] O. Szász, "Über die Approximation stetiger Funktionen durch lineare Aggregate von Potenzen," *Mathematische Annalen*, vol. 77 (1916), pp. 482-496.

[11] E. C. Titchmarsh, *The Theory of Functions*, 2nd ed., London, 1939.

UNIQUENESS THEOREMS FOR TWO FREE BOUNDARY PROBLEMS.*

By JAMES B. SERRIN, JR.

An important new approach to the Helmholtz problem of plane wakes, based upon general results in the theory of conformal representation, has been developed by Lavrentieff [1]. In his paper Lavrentieff considered the general existence and uniqueness for a certain class of obstacles; however, the essential simplicity and elegance of his method in the case of the uniqueness problem is obscured by his recourse to results of the complicated existence theory as well as by the difficulty one experiences in understanding his proof.[1]

It is the purpose of this paper to present an adequate and simple account of this uniqueness theory and to extend its validity. In Part I we develop the function theoretic basis for the theory. Theorem 1 is Lavrentieff's basic result, but the proof we give is new; on the other hand, Theorem 2 was neither stated nor proved by Lavrentieff, although he apparently made use of the result. In Part II the uniqueness theory proper is presented. In its exposition we have used, outside of Theorems 1 and 2, only one well known property of a flow solving the plane wake problem. The class of obstacles we consider includes as a special case the class considered by Lavrentieff, while it neither includes the class of obstacles discussed by Leray [2] nor is it included in that class.

Finally, in § 6 we show that flows solving the so-called finite cavity problem exist, while in § 7 we indicate certain generalizations.

Part I.

1. A corollary of Julia's theorem. Julia's theorem may be stated in the following form:[2] *Let* $f(\zeta)$ *be regular and* $|f(\zeta)| < 1$ *in* $|\zeta| < 1$. *Suppose*

* Received February 26, 1951. Prepared under Navy Contract N6onr-180, Task Order V, with Indiana University.

[1] The present author states frankly that he is unable to follow many of the arguments. Also see Weinstein, *Proc. Symposia on Applied Math.*, vol. 1, Am. Math. Soc., 1949, pp. 1-12.

[2] Bieberbach, Lehrbuch der Funktionentheorie, Chelsea, 1945, vol. 2, pp. 112-121.

that $f(1) = 1$ *and* $\lim f'(\zeta) = \alpha$, *where* $\zeta \to 1$. *Then* α *is real and positive, and*

$$(1) \qquad |1 - f(\zeta)|^2/(1 - |f(\zeta)|^2) \leq \alpha |1 - \zeta|^2/(1 - |\zeta|^2).$$

The equality holds if and only if $f(\zeta)$ *is linear.*

COROLLARY. *Let* $f(\zeta)$ *satisfy the hypotheses of Julia's theorem. In addition, suppose that* $f(-1) = -1$ *and* $\lim f'(\zeta) = \beta$, *where* $\zeta \to -1$. *Then* β *is real and positive, and* $\alpha\beta \geq 1$. *The equality holds if and only if*

$$(2) \qquad f(\zeta) = (\zeta + a)/(a\zeta + 1)$$

where a *is a real number,* $|a| < 1$.

Proof. The hypotheses of Julia's theorem are satisfied by the function $g(\zeta) = -f(-\zeta)$; in fact $\lim_{\zeta \to 1} g'(\zeta) = \lim_{\zeta \to 1} f'(-\zeta) = \beta$. Hence β is real and positive, and

$$(3) \qquad |1 + f(-\zeta)|^2/(1 - |f(-\zeta)|^2) \leq \beta |1 - \zeta|^2/(1 - |\zeta|^2).$$

Putting $\zeta = 0$ in (1) and (3), and multiplying the resulting inequalities together, we obtain

$$\alpha\beta \geq |1 - f(0)^2|^2/(1 - |f(0)|^2)^2 \geq 1.$$

Obviously, the equality can hold only if $f(\zeta)$ is linear. However, the only linear function satisfying the hypotheses is (2). For this function, $\alpha\beta = 1$, which completes the proof.

2. Applications to hydrodynamics. By a *flow* we shall mean a steady irrotational plane motion of an ideal incompressible fluid. We suppose that the motion occurs in a closed simply connected region R (which may be many sheeted) over the complex z-plane. The fluid may not penetrate the boundary of R (that is the boundary consists of streamlines), although a finite number of sources and sinks are allowed on the boundary. Let $w(z) = \phi + i\psi$ be the usual complex velocity potential; then $w(z)$ is regular in the interior of R and maps the boundary of R into straight arcs $\psi = \text{const.}$ of the w-plane.

Without exception, each flow we consider will occupy an infinite region and have a well defined velocity at infinity of fixed magnitude U.[3] A flow of this type will be called an *admissible flow* if $w(z)$ maps R onto the upper half plane $\psi \geq 0$. If in addition the inverse of $w(z)$ is single valued and

[3] By this we mean that the complex velocity $w'(z) = u + iv$ approaches a limit at infinity, and $|w'(\infty)| = U$.

regular for $\psi > 0$, then the flow will be called a *simple flow*. The speed at a point P in R of a flow in R will be denoted by $V(P, R)$; i. e.

$$V(P, R) \equiv | w'(z_P) |.^4$$

Since only one flow will be considered corresponding to a given region R, this notation will cause no confusion.

Let R_1 and R_2 be two regions of the type described above. If each point P of the surface R_2 may be set in correspondence with a point P' of R_1 such that $z_P = z_{P'}$, and if this correspondence is continuous for P on R_2, then we say that R_1 contains R_2. We shall often be dealing with several flows simultaneously; although these flows are considered as occurring in the same plane (z-plane), they are independent of one another. Our results will depend on the following theorems.

THEOREM 1. *Let R_1 and R_2 be two closed simply connected regions over the complex z-plane, such that R_1 contains R_2. Suppose that there exists an admissible flow in R_1 and a simple flow in R_2. If the streamlines $\psi = 0$ have a common point M, then $V(M, R_1) \geq V(M, R_2)$, the equality holding if and only if the flows are identical, or $V(M, R_1) = 0$.*

Proof. Let $w_1(z)$ and $w_2(z)$ denote the complex velocity potentials of the flows in R_1 and R_2 respectively. Without loss of generality we can take $w_1(z_M) = w_2(z_M) = 0$.

Define

$$\zeta(w) = \{-(i-1)w - (i+1)\}/\{(i-1)w - (i+1)\}.$$

This function maps the upper half plane, which we shall denote by Δ, onto the unit circle C such that $\zeta(0) = 1$, $\zeta(\infty) = -1$. If the unit circle C is mapped upon the upper half plane Δ by the inverse of $\zeta(w)$, Δ onto R_2 by the single valued regular function which is the inverse of $w_2(z)$, R_2 into a subregion of Δ by $w_1(z)$, and this subregion into C by $\zeta(w)$, we achieve a conformal mapping $f(\zeta)$ of the unit circle into itself. Clearly $f(1) = 1$, $f(-1) = -1$,

$$f'(\zeta) = \frac{\zeta'[w_1(z)] \cdot w_1'(z)}{\zeta'[w_2(z)] \cdot w_2'(z)} = \frac{w_1'(z)}{w_2'(z)} \left[\frac{(i-1)w_2(z) - (i+1)}{(i-1)w_1(z) - (i+1)}\right]^2.$$

---

[4] The complex number corresponding to a point P in the flow plane will be denoted by z_P. If P is on the boundary of R, $w'(z_P)$ does not necessarily exist. However, by writing the symbol $V(P, R)$ we shall imply that the derivative in question does exist.

Hence, recalling that α and β are real,

$$\alpha = \lim_{\zeta \to 1} |f'(\zeta)| = \lim_{z \to z_M} \left| \frac{w_1'(z)}{w_2'(z)} \right| \left| \frac{(i-1)w_2(z) - (i+1)}{(i-1)w_1(z) - (i+1)} \right|^2$$
$$= V(M, R_1)/V(M, R_2);$$
$$\beta = \lim_{\zeta \to -1} |f'(\zeta)| = \lim_{z \to \infty} |w_2(z)/w_1(z)|^2.$$

To evaluate the limit, we fix $\epsilon > 0$. For z sufficiently large, $|w_1'(z) - U| < \epsilon$, $|w_2'(z) - U| < \epsilon$. Using these inequalities, we obtain for sufficiently large z, z_A:

$$\frac{(U-\epsilon)|z - z_A| - |w_2(z_A)|}{(U+\epsilon)|z - z_A| + |w_1(z_A)|} \leq \left| \frac{w_2(z)}{w_1(z)} \right| \leq \frac{(U+\epsilon)|z - z_A| + |w_2(z_A)|}{(U-\epsilon)|z - z_A| - |w_1(z_A)|}.$$

Hence $(U-\epsilon)/(U+\epsilon) \leq \lim_{z \to \infty} |w_2(z)/w_1(z)| \leq (U+\epsilon)/(U-\epsilon)$. Since ϵ can be made arbitrarily small, the limit is 1. Thus, $\beta = 1$.

Applying the corollary, we get $\alpha\beta = V(M, R_1)/V(M, R_2) \geq 1$. The equality holds if and only if $f(\zeta) = (\zeta + a)/(a\zeta + 1)$; that is, if and only if the flows in R_1 and R_2 are identical.

The following lemma will be useful in proving Theorems 2 and 3.

LEMMA 1. *Let R_1 and R_2 be two regions whose boundaries have a common arc MN, and which lie on the same side of MN. Suppose $w_1(z)$ and $w_2(z)$, respectively, map R_1 and R_2 onto the upper half plane. If P is an interior point of MN and* $\lim |w_1'(z)|$, *where $z \to z_P$, exists, then* $\lim |w_2'(z)|$ *exists, where $z \to z_P$.*

The simple proof, which we omit, depends on the Schwarz reflection principle.

THEOREM 2. *Let R_1 and R_2 be two simply connected closed regions over the complex-z-plane. Suppose that there are admissable flows in R_1 and R_2, and let S_1 and S_2 denote the respective streamlines $\psi = 0$. Let S_1 and S_2 have an arc MN in common such that the direction of each flow on MN is from M to N. Furthermore let an arc QM of S_1 and an arc NQ of S_2 bound, together with MN, a simply connected region $MNQ \equiv R_3$ which is contained in both R_1 and R_2. Then the inequality*

$$V(M, R_2) < V(N, R_2), \tag{4}$$

implies that $V(M, R_1) < V(N, R_1)$.

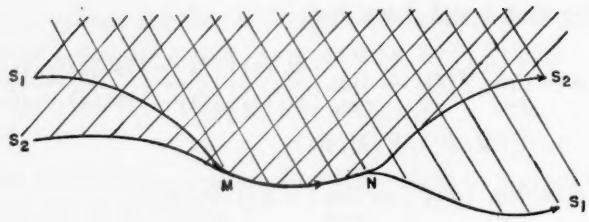

FIGURE 1

Note. Figure 1 shows a particularly simple case of the theorem when Q is the point at infinity and the flows in both R_1 and R_2 are simple. It is convenient to think of this situation rather than the more general one during the course of the proof, although the argument is identical in any case.[5]

Proof. Let $w_2(z)$ be the complex potential of the admissable flow in R_2. We may take $w_2(z_M) = 0$. Let $w_3(z)$ map R_3 conformally onto the upper half plane Δ such that $w_3(z_M) = 0$, $w_3(z_N) = w_2(z_N)$, $w_3(z_Q) = \infty$. By virtue of Lemma 1, $\lim_{z \to z_M} |w_3'(z)|$ and $\lim_{z \to z_N} |w_3'(z)|$ exist; we shall denote these limits by $V(M, R_3)$ and $V(N, R_3)$. The proof rests on two auxiliary inequalities.

First we shall show that

$$V(N, R_2)/V(N, R_3) \leq 1. \tag{5}$$

Let $\zeta(w)$ map Δ onto the upper half unit circle Γ such that $\zeta(0) = -1$, $\zeta[w_2(z_N)] = 0$, $\zeta(\infty) = +1$. The image of the positive real axis of the w-plane is then the segment $(-1, +1)$ of the real axis in the ζ-plane.

If Γ is mapped onto Δ by the inverse of $\zeta(w)$, Δ onto R_3 by the inverse of $w_3(z)$, R_3 into a subregion of Δ by $w_2(z)$, and this subregion into Γ by $\zeta(w)$, we achieve a mapping $f(\zeta)$ of Γ into itself. Now $f(0) = 0$, and $f(\zeta)$ is real and continuous on the real axis. Consequently, $f(\zeta)$ may be analytically continued, by reflection, throughout the entire unit circle. Then by Schwarz's Lemma, $f'(0) \leq 1$. But

$$f'(0) = \{\zeta'[w_2(z_N)]w_2'(z_N)\}/\{\zeta'[w_3(z_N)]w_3'(z_N)\} = V(N, R_2)/V(N, R_3).$$

Secondly, we shall show that

$$\{V(M, R_2)/V(M, R_3)\}\{V(N, R_2)/V(N, R_3)\} \geq 1. \tag{6}$$

---

[5] Theorem 2, as well as Theorem 1, is stated in more generality than is necessary for the purposes of this paper; in particular, the flows we consider in applications will all be simple. However, the increase in generality does not alter the proof, and in a later paper [3] it will be of the utmost importance.

It may be remarked here that if equality holds in (4), then equality holds in the conclusion if and only if the flows are identical.

Using a series of conformal mappings very similar to that already exhibited, we may define a function $g(\zeta)$ such that $|g(\zeta)| < 1$ for $|\zeta| < 1$, and $g(1) = 1$, $g(-1) = -1$, finally $g'(1) = V(M, R_2)/V(M, R_3)$, $g'(-1) = V(N, R_2)/V(N, R_3)$. The required result follows from our corollary.

Combining inequalities (5) and (6) we obtain

$$(7) \qquad V(N, R_2)/V(N, R_3) \leq 1 \leq V(N, R_3)/V(N, R_2) \leq V(M, R_2)/V(M, R_3).$$

Hence, using (7) and (4),

$$(8) \qquad V(M, R_3) \leq \{V(M, R_2)/V(N, R_2)\}V(N, R_3) < V(N, R_3).$$

Now a duplication of the preceding argument, this time with respect to the regions R_3 and R_1, leads to

$$(7') \qquad V(M, R_1)/V(M, R_3) \leq V(N, R_1)/V(N, R_3);$$

(it is sufficient to interchange the roles of M and N, and to replace R_2 by R_1). From (7′) and (8) we obtain

$$V(M, R_1) \leq \{V(M, R_3)/V(N, R_3)\}V(N, R_1) < V(N, R_1),$$

which completes the proof.

The following theorem is essentially due to Lavrentieff. Although no use will be made of its result, it is an interesting corollary of the preceding theorems. Let R be a region occupied by a simple flow, and let S be the streamline $\psi = 0$. We may assume without loss of generality that S has zero tangent at infinity and that R is "above" S. Let S contain an arc M^*PN^* having the following properties:

1. If S is given parametrically by $z = z(\phi)$, then the points $-\infty$, M^*, P, N^*, $+\infty$ correspond in that order to increasing ϕ;

2. M^*PN^* has a continuous tangent;

3. P is an inclination point of S, and the inclination of S is monotone on M^*P and PN^*;

4. The straight line L tangent to S at P does not intersect S other than at P, and is not horizontal.

5. The speed $V(Q, R)$ exists at all interior points Q of M^*PN^*.

A direction is assigned to L to correspond to the direction of increasing ϕ at P; Figure 2.

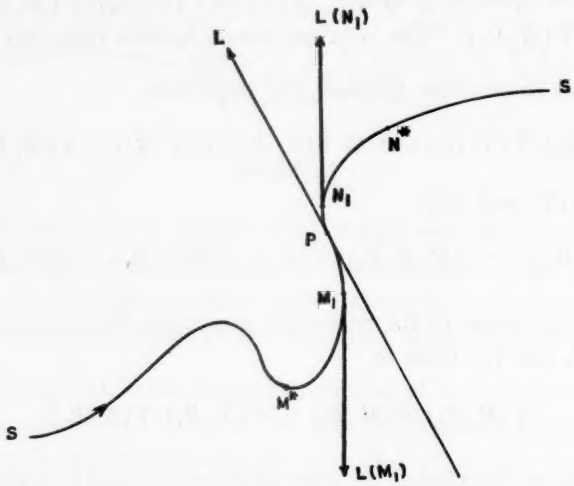

FIGURE 2

THEOREM 3. *Under hypotheses* 1-5, *there is an arc* MPN *of* M^*PN^* *such that, if* L *is directed into* R, $V(N,R) > V(M,R)$, *or if* L *is directed out of* R, $V(N,R) < V(M,R)$.

Proof. We note here that no attempt has been made to obtain a minimal list of hypotheses for the validity of the theorem, or even to state the present hypotheses in the shortest possible form. It is sufficient to consider only the case that L is directed into R.

Let N_0 be any point of PN^*, and denote by $L(N_0)$ the ray tangent to S at N_0, directed in the sense of increasing ϕ at N_0. Similarly let M_0 be any point of M^*P, and denote by $L(M_0)$ the ray tangent to S at M_0, directed in the sense of decreasing ϕ at M_0.

In virtue of hypotheses 1-4 there exists an arc M_1PN_1 of M^*PN^* such that $L(N_1)$ is entirely in R, $L(M_1)$ is entirely outside R, $L(N_1)$ and $L(M_1)$ have the same inclination, and the inclination of S on both M_1P and PN_1 is monotone. See figure 2.

The simple closed curve consisting of the ray $L(M_1)$, the arc M_1PN_1, and the ray $L(N_1)$ divides the plane into two simply connected regions; let R_2 be that one containing the arc $(-\infty, M_1)$ of S. By Lemma 1 we deduce easily that R_2 contains a simple flow. Moreover, by the same lemma, $V(Q,R_2)$

exists for interior points Q of M_1PN_1. We shall prove that there is a subarc MPN of M_1PN_1, such that

$$V(N, R_2) > V(M, R_2). \tag{9}$$

Assuming for the moment that this is true, then the hypotheses of Theorem 2 are satisfied by R and R_2 (compare figures 1 and 2). It follows that $V(N, R) > V(M, R)$, proving the theorem.

Before proving (9) we shall introduce a useful notation: any region which is congruent to a given region R will be denoted by R'. If a point in R', when considered as a geometrical entity of the configuration R', corresponds under the congruence to a point Q of R, it will be denoted by Q', etc.

Now we can place (or more precisely, construct) a replica R_2' of R_2 such that R_2' is contained in R_2 and some interior point M' of $M_1'P'$ coincides with some interior point N of PN_1. Then by Theorem 1, $V(N, R_2) > V(M', R_2') = V(M, R_2)$, proving (9).

Part II.

3. The problem of the wake. Consider a fixed obstacle (arc) AB immersed in a uniform flow (§ 2) having a velocity U at infinity. It is assumed that the flow separates from AB at its endpoints, leaving a *wake* or *cavity* behind AB. The streamlines Σ and Σ^* which bound the cavity are not given geometrically, but by the condition that the flow speed on Σ and Σ^* is to be constant; this constant will be called the *cavity speed*. The *problem of the wake* is to determine such a flow past the given obstacle AB (a fuller discussion of this well known problem is readily available in nearly every treatise on hydrodynamics). We shall consider the uniqueness problem for two characteristically different wakes that an obstacle might produce.

First we consider the usual *Helmholtz wake problem.* That is, Σ and Σ^* are required to extend downstream to infinity; it is well known that this can occur only when the cavity speed is U.

The second problem is the so-called *finite cavity problem.*[6] As the name implies, it is required that the cavity be finite in extent, so that i) the free streamlines meet and end at a point E behind the obstacle, and ii) the point at infinity is an interior point of the flow.

Let the flow take place in the complex z-plane. There is obviously a stagnation point O of the flow on AB; O will be taken as origin and the positive x-axis in the direction of the uniform flow. For the purposes of this

[6] A fuller discussion of the finite cavity problem is contained in [3].

paper *it is necessary to restrict our consideration henceforth to symmetric flows past symmetric obstacles.* Because of this we need only look at the "upper half" of the complete flow and, indeed, this point of view will be adopted consistently. We shall assume that AB has a continuous tangent; the inclination of the tangent vector will be denoted by the variable Ψ.[7] A symmetric obstacle will be called a *regular obstacle* if there exists a number α, $\pi/2 \leq \alpha \leq \pi$, such that $\alpha - \pi \leq \Psi \leq \alpha$ holds along OB.

Regular obstacles will be our main interest. Clearly a line whose inclination is α intersects the arc OB of a regular obstacle in at most a single point or along a straight segment. The following lemma shows that Σ enjoys the same property.

LEMMA 2. *In a symmetric solution of either the Helmholtz problem or the finite cavity problem for a regular obstacle, the inclination θ of Σ satisfies $\alpha - \pi < \theta < \alpha$.*

A proof of Lemma 2 is given in [3]. The following result is well known.

LEMMA 3. *The free streamline Σ in a symmetric solution of the Helmholtz problem may be represented parametrically by the equations*

$$x = x(\zeta) = a[\zeta^{-2} + \omega'(0)^2 \log |\zeta| + P_1(\zeta^2)],$$

$$y = y(\zeta) = a[2\omega'(0)\zeta^{-1} + P_2(\zeta^2)],$$

where $-1 < \zeta \leq 0$. Here a and $\omega'(0)$ are certain constants, and $P_1(\zeta)$ and $P_2(\zeta)$ are analytic functions in the unit circle. Hence the free streamlines are analytic curves.

Remark. If $\omega'(0) \neq 0$, the free streamlines approach the parabola $y^2 = [4a\omega'(0)^2]x$. If $\omega'(0) < 0$, Σ approaches the upper branch, and Σ^* the lower. If $\omega'(0) > 0$, Σ approaches the lower branch, and Σ^* the upper. This overlapping of the flow region in the latter case, although physically unrealistic, is not objectionable mathematically. If $\omega'(0) = 0$, then the free streamlines asymptotically approach a curve $y^2 = \text{const.}\, x^{-2n-1}$, where n is a positive integer.

In the finite cavity problem a corresponding lemma exists, and the free streamlines are analytic curves. We note finally that the "upper half" of a flow solving either of our problems is a simple flow. This fact is important in the application of Theorems 1 and 2.

---

[7] In determining Ψ, the tangent vector is considered to be directed from A to B; inclination is measured in a counterclockwise sense from the positive x-axis.

4. The Helmholtz problem. We recall that it is only necessary to consider the "upper half" of a flow.

THEOREM 4. *There can be no more than one symmetric flow solving the Helmholtz problem for a regular obstacle AB.*

Proof. Suppose for contradiction that there exist two flows solving this problem. Let R_1 and R_2 be the regions occupied by these flows, and let Σ_1 and Σ_2 be the corresponding free streamlines. It is apparent by Lemma 3 that one of the free streamlines, say Σ_1, is below the other at infinity.

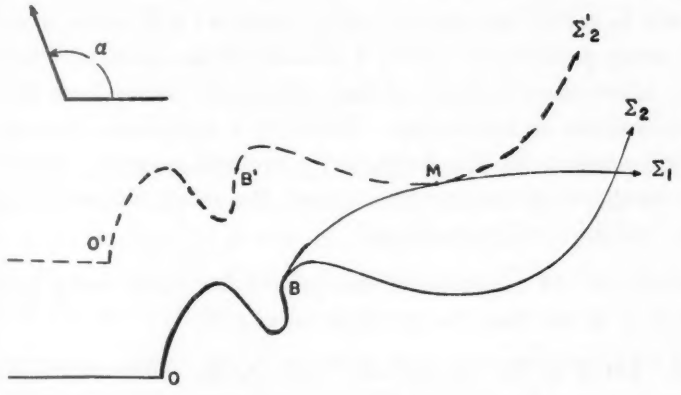

FIGURE 3

Figure 3 illustrates the situation arising from a suitable translation of a replica R_2' of R_2. (In figures, we represent a replica of a given region by a dotted curve. The prime notation associated with a replica is explained on page ——.) Here R_1 contains R_2', and Σ_1 and Σ_2' have a common point M. It will be proved in a moment that such a translation always exists (provided only that AB is a regular obstacle). Assuming this result, however, we obtain an immediate contradiction. For on one hand, by Theorem 1, $V(M, R_1) > V(M, R_2')$, while on the other hand, $V(M, R_1) = U = V(M, R_2')$, since M is on both the free streamlines Σ_1 and Σ_2'.

To complete the proof, we begin by noting a property of the curves Σ_1 and Σ_2 which follows from Lemma 2. Namely, let Λ be the line passing through B whose inclination is α. Then any line parallel to Λ,

i) cannot intersect either Σ_1 or Σ_2 if it is on the same side of Λ as OB;

ii) intersects both Σ_1 and Σ_2 in exactly one point if it is on the other side of Λ.

Now let $z = z(s)$ be a parametric representation of Σ_2. Let $d(s)$ be the distance between the point $z(s)$ of Σ_2 and the curve Σ_1, measured along a line parallel to Λ. (We take $d(s)$ positive if the distance is measured into R_2 and negative otherwise.) In virtue of ii), the function $d(s)$ is well defined, and by Lemma 3 it is continuous. $d(s)$ is zero at B and negative at infinity; therefore, it attains a maximum value d.

If the replica R_2', originally in coincidence with R_2, is translated a distance d parallel to Λ, the required situation is evidently obtained, q. e. d.

5. The schlicht cavity problem. Given an obstacle AB, the following problem will be called the *schlicht cavity problem*: find *either* a solution of the finite cavity problem for AB *or* a solution of the Helmholtz problem for AB; in the latter case it is required that, sufficiently distant from the obstacle, the free streamlines do not overlap. Evidently a uniqueness theorem for this problem will account for the finite cavity problem as well. Before stating our main result, we prove two preliminary theorems, the second justifying the name "schlicht cavity problem."

THEOREM 5. *In a symmetric flow solving the finite cavity problem the cavity speed V is less than the speed at infinity U.*

Proof. Let R be the "upper half" flow region. Obviously there is some point M on Σ with a non-positive minimal ordinate y_M. Let R_1 be half plane above the line $y = y_M$. The uniform flow $w(z) = Uz$ in R_1 is a simple flow, and R_1 contains R. Therefore from Theorem 1,

$$V = V(M, R) < V(M, R_1) = U.$$

THEOREM 6. *In a symmetric flow solving schlicht cavity problem the free streamline Σ stays entirely above the real axis.*[8]

Proof. Let R be the "upper half" flow region. We prove separately case I, when the flow solves the Helmholtz problem, and case II, when the flow solves the finite cavity problem.

Case I. Suppose the theorem is false. Then there must be some point M on Σ with a non-positive minimal ordinate y_M. Let R_1 be the half plane above the line $y = y_M$. Then by Theorem 1, $V(M, R_1) > V(M, R)$. But this is impossible, for $V(M, R) = U = V(M, R_1)$.

Case II. Let M_1 be the point of Σ with minimal ordinate, and let N_1

[8] Except, of course, at E if the flow solves the finite cavity problem.

be the point of the arc M_1E of Σ with maximal ordinate. Let R_2 be the region above the curve consisting of the arcs $L(M_1)$, M_1N_1, and $L(N_1)$ in figure 4.

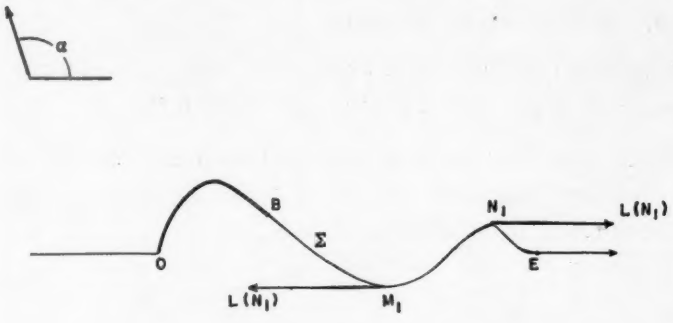

FIGURE 4

Now by Theorem 1, there exist two points M and N of the arc M_1N_1 such that $V(M, R_2) < V(N, R_2)$; (see the proof of Theorem 3). The hypotheses of Theorem 2 are satisfied by R and R_2 (compare figures 1 and 4). It follows that $V(M, R) < V(N, R)$. This is impossible since $V(M, R) = V(N, R) = V$, where V is the cavity speed.

THEOREM 7. *There can be no more than one symmetric flow solving the schlicht cavity problem for a given regular obstacle AB.*

Proof. Suppose for contradiction that there exist two flows solving this problem. Let R_1 and R_2 be the regions occupied by these flows, and let Σ_1 and Σ_2 be the corresponding free streamlines. If both flows solve the Helmholtz problem we contradict Theorem 4. There remain two cases.

Case I. The flow in R_1 solves the finite cavity problem, while the flow in R_2 solves the Helmholtz problem. Figure 5 illustrates the situation arising

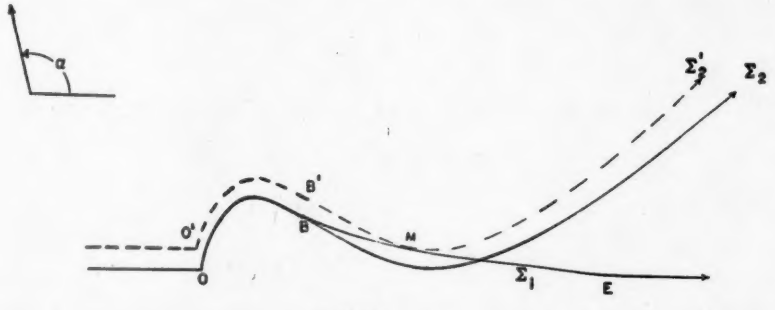

FIGURE 5

from a suitable translation of a replica R_2' of R_2. Here R_1 contains R_2', and Σ_1 and Σ_2' have a common point M. A rigorous justification of this assertion is readily obtained by the same reasoning already employed in the proof of Theorem 4, and need not be repeated.

By Theorem 1, $V(M, R_1) > V(M, R_2')$. This is impossible, however, for (Theorem 5) $V(M, R_1) = V < U$ and $V(M, R_2') = U$.

Case II. Both the flow in R_1 and the flow in R_2 solve the finite cavity problem. We shall suppose the endpoint E_1 of the streamline Σ_1 is to the right of the endpoint E_2 of the streamline Σ_2.

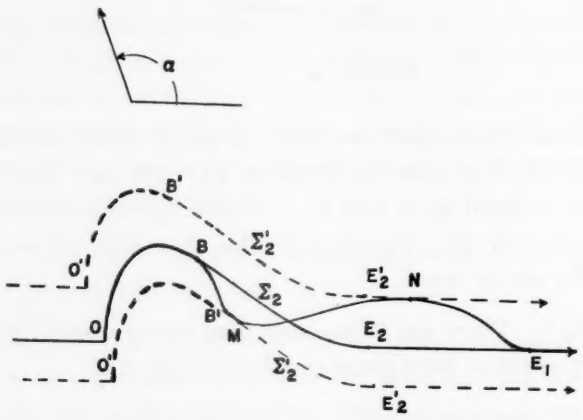

FIGURE 6

Figure 6 illustrates the situation arising from *two* suitable translations of replicas R_2' of R_2. Consider first the replica R_2' which contains R_1. Here, $V(M, R_2') > V(M, R_1)$. However, since M is on both the free streamlines Σ_2' and Σ_1, we have $V(M, R_2') = V_2$ and $V(M, R_1) = V_1$, where V_1 and V_2 are the cavity speeds. Hence

$$V_2 > V_1. \tag{10}$$

On the other hand, considering the replica R_2' inside R_1, we have

$$V(N, R_1) > V(N, R_2'). \tag{11}$$

Now N is on Σ_1, so that

$$V(N, R_1) = V_1. \tag{12}$$

However, considered as a point of R_2', N may be on either the free streamline

Σ_2' or the arc (E_2', ∞). (The latter possibility is shown in figure 6.) If N is on Σ_2' then

$$V(N, R_2') = V_2. \tag{13}$$

From (11)-(13) it follows that $V_1 > V_2$. This contradicts (10), so that N must be on (E_2', ∞). If we can show that in this case

$$V(N, R_2') > V_2, \tag{14}$$

then it will follow again that $V_1 > V_2$, contradicting (10) and completing the proof.

To prove (14), it is sufficient to show that $V(N, R_2) > V_2$, for any point N on the arc (E_2, ∞). Thus, let the region R_3 be the upper half plane. If the common arc EN of the boundaries of R_2 and R_3 is associated with the arc MN of Theorem 2, then the geometrical hypotheses of Theorem 2 are satisfied by R_2 and R_3. Moreover,

$$V(E, R_3) = V(N, R_3) = U. \tag{4'}$$

Hence $V_2 = V(E, R_2) < V(N, R_2)$. q. e. d.

6. Existence of finite cavity flows. Consider a simple closed curve T symmetric about the x-axis. Let O and O' be the two points of intersection of T with the x-axis, with O to the left of O'. Let T have the property that any symmetric arc BOB^* of T is a regular obstacle.

It can be shown [3] with the help of the Schauder-Leray fixed point theory that there exists a solution of the schlicht cavity problem for any symmetric arc BOB^* of T. If B, B^* are sufficiently near O, then any line parallel to the x-axis can intersect BOB^* in at most one point. In this case it is well known (see Leray [2], p. 157) that the (unique) solution which does exist is necessarily a Helmholtz flow. Conversely, we can guarantee that if B and B^* are sufficiently near O' the solution which does exist is necessarily a finite cavity.

Proof. Consider first the symmetric classical flow, with uniform velocity U at infinity, past the obstacle T. The real axis exterior to T is a streamline of this flow, while points O and O' are stagnation points. Since O' is a stagnation point there is some arc $NO'N^*$ of T along which the speed is everywhere less than or equal to U. We assert that if BOB^* is a symmetric arc of T and B, B^* are points of $NO'N^*$, then the solution of the schlicht cavity problem cannot be the Helmholtz flow. For, upon supposing the contrary, essentially the argument used in the proof of Theorem 7, Case I, leads to an impossibility.

7. Concluding remarks. The results of this paper carry over as well to symmetric flows past obstacles with corners. More important is the following generalization.

A curve will be called *starlike* if

i) there exists a point σ in the lower half plane $y \leq 0$, such that any straight line passing through σ intersects the curve in at most one point or along a segment, or if

ii) there exists a number α, $0 \leq \alpha \leq \pi$, such that any straight line whose inclination is α intersects the curve in at most one point or along a segment.

A symmetric obstacle with a continuous tangent (a finite number of corners may be allowed) will be called a *starlike obstacle* if the curve consisting of the negative real axis together with OB is starlike. It is apparent that the regular obstacles discussed in the body of the paper are also starlike obstacles. The proof of the following lemma is indicated in [3].

LEMMA 2′. *In a symmetric solution of the schlicht cavity problem for a starlike obstacle AB, the steramline $\psi = 0$ consisting of the negative real axis, OB, and Σ is a starlike curve.*

With the help of this lemma, Theorems 5-7 may easily be shown to hold for starlike obstacles. The proof need be modified only to the extent that our auxiliary translations of replicas of given flows regions be replaced by suitable magnification of replicas, the center of magnification being σ. Such a magnification, it is noted, leaves flow velocities unchanged at geometrically similar points.

UNIVERSITY OF INDIANA.

REFERENCES.

[1] M. Lavrentieff, "Sur certaines proprietes des fonctions univalent et leurs applications a la theorie des sillages" (in Russian), *Recueil Mathematique (Matematicheskii Sbornik)*, vol. 46 (1938), pp. 391-458.

[2] J. Leray, "Les problemes de representation conforme d'Helmholtz; theorie des sillages et des proues," *Commentarii Mathematici Helvetici*, vol. 8 (1935-36), pp. 149 and 250.

[3] J. Serrin, "Existence theorems for some free boundary problems," *Journal of Rational Mechanics and Analysis*, vol. 1 (1952), pp. 1-48.

EXTENSIONS OF DIFFERENCE FIELDS.*

By RICHARD M. COHN.[1]

1. Introduction. By an (abstract) difference field [4][2] we mean a field with an endomorphism. The algebraic theory of difference equations has led to the discovery of two surprising phenomena concerning difference fields. J. F. Ritt [5], § 3, showed that a difference field $\mathfrak{F}$ may have extensions $\mathfrak{G}$ and $\mathfrak{H}$ which cannot both be embedded in any one extension of $\mathfrak{F}$, and which we shall therefore call incompatible. In [1] it was shown that a difference field $\mathfrak{F}$ of characteristic 0 may have a proper extension which does not permit two distinct mappings into any extension of $\mathfrak{F}$. We call such an extension a monadic extension of $\mathfrak{F}$. A corresponding phenomenon occurs in the theory of algebraic fields, that is, fields in the ordinary sense, only if the characteristic p exceed 0. In this case an extension which is monadic in the sense just described may be produced by adjoining a p-th root not already in the field.

Our purpose in this paper is to present in organized form the still rudimentary theory of these phenomena, and to point out their decisive importance in the algebraic theory of difference equations. They seem, in fact, to mark a point byond which one can no longer use the theory of polynomial ideals or the algebraic theory of differential equations as a guide to the study of difference equations, but must expect phenomena which are *sui generis*. Our results also suggest interesting questions concerning field structure. It would, for example, be desirable to characterize those fields which possess incompatible or monadic extensions. Our principal result is a first step toward such a characterization; *fields which are algebraically closed have no finite extensions of these types.*

2. Definitions. By the (first) transform e_1 of an element e of a difference field $\mathfrak{F}$ we mean the element into which e is carried by the endomorphism in virtue of which $\mathfrak{F}$ is a difference field. By the transform of order k, or the k-th transform e_k of e we mean, of course, the element into which e is carried by the k-th iterate of the endomorphism (hence $e_0 = e$).

* Received April 28, 1951.

[1] This paper was written while the author held a university research grant from Rutgers University.

[2] Numbers in square brackets refer to papers listed in the Bibliography.

For the difference field $\mathfrak{G}$ to be an extension of the difference field $\mathfrak{F}$, it must be an extension of $\mathfrak{F}$ as an algebraic field, and the operation of taking transforms in $\mathfrak{G}$ must coincide for elements in $\mathfrak{F}$ with that of taking transforms in $\mathfrak{F}$. The difference field resulting from the adjunction of a set S of elements of $\mathfrak{G}$ to $\mathfrak{F}$ is by definition the least extension of $\mathfrak{F}$ containing the elements of S and their transforms of all orders. Such an extension will be called a finite (countable, etc.) extension of the difference field $\mathfrak{F}$ if S is finite (countable, etc.). Hence a finite extension of the difference field $\mathfrak{F}$ may well be an infinite extension of $\mathfrak{F}$ considered as an algebraic field. The extension just described will be denoted by $\mathfrak{F}\langle S\rangle$, while the algebraic field resulting from the adjunction of the elements of S without their transforms will be denoted by $\mathfrak{F}(S)$.

We define two extensions $\mathfrak{G}$ and $\mathfrak{H}$ of a difference field $\mathfrak{F}$ to be *incompatible* if there does not exist an extension $\mathfrak{K}$ of $\mathfrak{F}$ containing a subfield $\mathfrak{G}_1$ isomorphic to $\mathfrak{G}$, and a subfield $\mathfrak{H}_1$ isomorphic to $\mathfrak{H}$, these isomorphisms leaving each element of $\mathfrak{F}$ fixed.[3]

Let $\mathfrak{G}$ be an extension of $\mathfrak{F}$ such that not every element of $\mathfrak{G}$ has a transform of some order in $\mathfrak{F}$. We then define $\mathfrak{G}$ to be a *monadic extension* of the difference field $\mathfrak{F}$, if there does not exist an extension $\mathfrak{K}$ of $\mathfrak{F}$ into which $\mathfrak{G}$ has two distinct isomorphisms both preserving the elements of $\mathfrak{F}$.[4]

All fields occurring henceforth in this paper will be supposed to be of characteristic 0.

3. Examples and applications. *Example* 1. Let $\mathfrak{F}$ be the difference field consisting of the real numbers, with the identity as endomorphism. Let $\mathfrak{G}$ be the difference field consisting of the complex numbers, with the identity as endomorphism, and let $\mathfrak{H}$ be the difference field consisting of the complex numbers, with the transform of a number defined to be its complex conjugate. Since an extension of $\mathfrak{F}$ can contain at most one subfield isomorphic as an algebraic field to the field of complex numbers, it is evident that $\mathfrak{G}$ and $\mathfrak{H}$ are incompatible.

Let us consider the systems of difference polynomials[5] with coefficients in $\mathfrak{F}$

---

[3] Henceforth, when we refer to isomorphisms of extensions of a difference field $\mathfrak{F}$, it will be understood that we mean isomorphisms which leave each element of $\mathfrak{F}$ fixed.

[4] In other words no extension of $\mathfrak{F}$ can contain two distinct subfields each isomorphic to $\mathfrak{G}$ by isomorphisms which leave every element of $\mathfrak{F}$ fixed, nor does $\mathfrak{G}$ have an automorphism, other than the identity, which leaves each element of $\mathfrak{F}$ fixed.

[5] We deal here with a difference ring $\mathfrak{F}\{y\}$ consisting of all polynomials in

A) $y^2 + 1,\ y_1 - y,$ and B) $z^2 + 1,\ z_1 + z.$

It is evident that A) has a solution, the square root of -1, which generates $\mathfrak{G}$ when adjoined to $\mathfrak{F}$, while B) has a solution, again the square root of -1, which generates $\mathfrak{H}$. Hence it is possible to solve the system A) or the system B), but not to solve these systems simultaneously, even though they have no unknown in common.

We may put the matter in a more basic, if less striking, form by saying that the manifold of the algebraically irreducible difference polynomial $y^2 + 1$ consists of two irreducible manifolds [6] $\mathfrak{M}_1$ and $\mathfrak{M}_2$. $\mathfrak{M}_1$ is the manifold of the reflexive prime difference ideal $\mathfrak{Z}_1$ with characteristic set [7] $y^2 + 1,\ y_1 - y,$ and $\mathfrak{M}_2$ is the manifold of the reflexive prime difference ideal $\mathfrak{Z}_2$ with characteristic set $y^2 + 1,\ y_1 + y$. No extension of $\mathfrak{F}$ contains generic zeros [7] or, indeed, solutions of both $\mathfrak{Z}_1$ and $\mathfrak{Z}_2$. As suggested by this example, we define a set of reflexive prime difference ideals (or their manifolds) to be *incompatible* if there exists no extension of the coefficient field containing generic zeros of all of them. We say that an algebraically irreducible difference polynomial has an *inconsistent general solution*, if the manifolds of its general solution (see [1], p. 152) are incompatible. We shall later see that if a

---

$y_0 = y, y_1, y_2 \cdots$, with coefficients in $\mathfrak{F}$. The endomorphism of the ring is defined by the stipulations that elements of $\mathfrak{F}$ shall have the same transforms as in $\mathfrak{F}$ itself, and that the transform of y_i shall be y_{i+1}, $i = 0, 1, \cdots$. To solve a system of polynomials of $\mathfrak{F}\{y\}$, we must find a difference field $\mathfrak{G}$ which is an extension of $\mathfrak{F}$, and contains an element a such that when y is replaced by a (which means, of course, that, for each i, y_i is replaced by the i-th transform of a), the polynomials vanish. The extension of these concepts to rings in two indeterminates y, z, or rings in n indeterminates, is obvious. In the latter case we shall generally designate the indeterminates by $y_1, y_2, \cdots, y_n$. Then y_{ij} will denote the j-th transform of y_i, $i = 1, \cdots, n$; $j = 0, 1, \cdots$.

[6] The manifold of a system of difference polynomials is the set of its solutions in all extensions of the coefficient field. It is proved in [3] that every such manifold of solutions is the union of a finite number of irreducible manifolds, that is of manifolds which, if they annul a product AB of difference polynomials, annul one of the factors. The statement in [3] does not refer to abstract difference fields, but the work of [4] makes it evident that the result applies to them also. The set of all polynomials vanishing on an irreducible manifold forms a reflexive prime difference ideal. Conversely, the solutions of a reflexive prime difference ideal constitute an irreducible manifold. This correspondence is one-to-one.

[7] Characteristic sets of difference polynomials are defined in [3], where they are called basic sets. In this paper we have brought the terminology for difference polynomials into accord with the new terminology for differential polynomials adopted by Ritt in [6]. We list here the terms used in earlier papers, each followed by the term which is to replace it: arbitrary unknown, parametric indeterminate; ascending set, chain; basic sequence, characteristic sequence; basic set, characteristic set; general point, generic zero.

difference field has a finite extension which is incompatible with some other extension, then there is an algebraically irreducible zero-order difference polynomial with coefficients in the field, whose general solution is inconsistent.

As a result of the existence of incompatible extensions, the dimension or order [8] of a reflexive prime difference ideal may be reduced when the coefficient field is extended. Thus the system A), which is of zero order in $\mathfrak{F}$, has no solution if $\mathfrak{H}$ is taken as the coefficient field. The system

C) $$y_k^2 + y^2, \qquad y_{k+1}y - y_k y_1,$$

is the characteristic set of a prime ideal Σ of order k with coefficients in $\mathfrak{F}$, as we know from Theorem IV of [1]. If α is any solution of Σ except 0, then since α satisfies C), α_k/α is a solution of A). Hence $\mathfrak{F}\langle\alpha\rangle$ is an extension of $\mathfrak{G}$. It follows that 0 is the only solution [9] of the polynomials of Σ, considered as a system of polynomials with coefficients in $\mathfrak{H}$. Hence the order of this system is reduced from k to 0 when the coefficient field is extended to $\mathfrak{H}$.

Incompatible extensions also lead to difficulties when we try to replace parametric indeterminates by special values. From the analogy with systems of algebraic or algebraic differential equations, we should expect that if Π is a reflexive prime difference ideal in the indeterminates $u_1, u_2, \cdots, u_q$; $y_1, y_2, \cdots, y_p$, the u_i being parametric indeterminates,[10] it would be possible to find a solution of Π in which the u_i have any assigned values not annulling a certain difference polynomial U in the u_i. But this is not true for the reflexive prime ideal Π in indeterminates [11] u, y, with characteristic set

D) $$u^2y^2 + 1, \qquad u_1y_1 + uy.$$

Here u is a set of parametric indeterminates for Π, and the coefficient field is $\mathfrak{F}$ above. Suppose a U of order m exists. Let α be a generic zero of a system C) with $k > m$. Then Theorem IV of [1] assures us that α does

---

[8] "Dimension" corresponds to the heuristic concept of the number of indeterminates whose values can be chosen arbitrarily (parametric indeterminates). "Order" corresponds to the heuristic concept of the number of arbitrary periodic functions in the solution. See [1], p. 141 and p. 162.

[9] That 0 is a solution of Σ can be shown by an easy application of the technique used in our paper "Singular manifolds of difference polynomials," *Annals of Mathematics*, vol. 53, pp. 445-463. Let ϕ be a generic zero of Σ. Adjoin to $\mathfrak{F}$ an element c which is its own transform, and which satisfies no algebraic equation over $\mathfrak{F}$. Then ϕc annuls the polynomials of C) but not their initials. Hence ϕc annuls the polynomials of Σ. It follows readily from this that 0 is in the manifold of Σ.

[10] Here single subscripts do not indicate transforms.

[11] Single subscripts do indicate transforms here.

not annul U. Now $\mathfrak{F}\langle\alpha\rangle$ is an extension of $\mathfrak{G}$, while every solution of Π lies in an extension of $\mathfrak{H}$. Hence there cannot be a solution of Π with $u = \alpha$. Thus the assumption that U exists leads to a contradiction.

The existence of incompatible extensions also has implications for the analytic theory of difference equations. We consider difference polynomials in a field $\mathfrak{Z}$ of functions of the complex variable z meromorphic on a line $\mathfrak{L}$ which contains $z+1$ if it contains z. The transform of a function $f(z)$ is defined to be $f(z+1)$. If S_1 and S_2 are systems of such difference polynomials with the property that no extension of $\mathfrak{Z}$ as an abstract difference field contains both a solution of S_1 and a solution of S_2, then obviously at most one of these systems has solutions meromorphic on $\mathfrak{L}$. Of the systems A) and B) the latter has, of course, no meromorphic solutions whatsoever.

Example 2. We consider the field of rational functions of x with complex coefficients. This becomes a difference field $\mathfrak{F}$, if we define the transform of any rational function $f(x)$ to be the function $f(4x)$. Consider the difference polynomial $A = y^2 - x$ which is algebraically irreducible in $\mathfrak{F}$. Its transform A_1 is $y_1{}^2 - 4x$. Then $A_1 - 4A = y_1{}^2 - 4y^2 = (y_1 - 2y)(y_1 + 2y)$. Hence every solution of A annuls either

E) $y^2 - x, y_1 - 2y$ or F) $y^2 - x, y_1 + 2y$.

We know from Theorem IV of [1] that both E) and F) have solutions in suitable extensions of $\mathfrak{F}$. We shall show that no one extension contains solutions of both, so that A has an inconsistent general solution.

Let α be a solution of E). The only other solution of A lying in an extension of $\mathfrak{F}\langle\alpha\rangle$ is $\beta = -\alpha$. From E) we have $\alpha_1 = 2\alpha$. Now $\beta_1 = -\alpha_1 = -2\alpha = 2\beta$. Hence β annuls E), and no extension of $\mathfrak{F}\langle\alpha\rangle$ contains a solution of F).

Example 3. We now define the transform of a rational function $f(x)$ with complex coefficients to be $f(x^3)$, and let $\mathfrak{F}$ be the difference field consisting of all such functions of x with transforms defined in this way. We find that every solution of the algebraically irreducible difference polynomial $y^2 - x$ annuls either

G) $y^2 - x, y_1 - yx$ or H) $y^2 - x, y_1 + yx$.

As before we see that the manifolds of G) and of H) are incompatible.

Example 4. If we define the transform of a rational function $f(x)$ with complex coefficients to be $f(x^2)$, we get a difference field $\mathfrak{F}$ which has monadic

extensions. For example, every solution of the algebraically irreducible difference polynomial $y^3 - x$ annuls one of the systems

I) $y^3 - x, y_1 - y^2,$ J) $y^3 - x, y_1 - \omega y^2,$

K) $y^3 - x, y_1 - \omega^2 y^2,$

where ω is a complex cube root of unity. All these systems have solutions in suitable extensions of $\mathfrak{F}$. Let α be a solution, say of I). Then $\omega\alpha$, $\omega^2\alpha$, are the only other solutions of $y^3 - x$ in any extension of $\mathfrak{F}\langle\alpha\rangle$. It is easy to see that they annul the systems K) and J) respectively. Hence no extension of $\mathfrak{F}\langle\alpha\rangle$ contains a subfield other than $\mathfrak{F}\langle\alpha\rangle$ itself which is isomorphic to $\mathfrak{F}\langle\alpha\rangle$ by an isomorphism leaving each element of $\mathfrak{F}$ fixed; nor does $\mathfrak{F}\langle\alpha\rangle$ have any automorphism except the identity which leaves each element of $\mathfrak{F}$ fixed. No transform of α is in $\mathfrak{F}$. Hence $\mathfrak{F}\langle\alpha\rangle$ is a monadic extension of $\mathfrak{F}$. J) and K) likewise furnish monadic extensions of $\mathfrak{F}$.

It is interesting to note that if we define the transform of a rational function $f(x)$ with complex coefficients to be $\bar{f}(x^2)$, where $\bar{f}$ is the function obtained by replacing each coefficient in f by its complex conjugate, we obtain a difference field $\mathfrak{G}$ in which the systems I), J), K) define not monadic extensions, but rather extensions any two of which are incompatible.

We define a reflexive prime difference ideal $\mathfrak{Z}$, not containing 1, or its manifold $\mathfrak{M}$ to be monadic, if no transform of a generic zero of $\mathfrak{Z}$ lies in the coefficient field, and if there exists no extension of the coefficient field containing two generic zeros of $\mathfrak{Z}$. We shall see that if a difference field $\mathfrak{F}$ has finite monadic extensions (and consequently monadic prime ideals), it has a monadic extension which is obtained by adjoining a solution of a zero order difference polynomial in one indeterminate. In fact, the situation with finite monadic extensions is even stronger than with finite incompatible extensions—every finite monadic extension is of zero effective order.[12]

The proof that finite algebraic extensions of a field of characteristic zero are simple, and the proof of the corresponding theorem for differential fields,[13] are based on the fact that such fields do not have monadic extensions. For fields of characteristic p the proof fails and the theorem does not hold. In the case of difference fields of characteristic zero, however, the difficulty

---

[12] The concept of a quasi-linear system, defined in [1], p. 164, includes monadic prime ideals and the uninteresting prime ideals whose generic zeros or their transforms lie in the coefficient field. The result we have just referred to is essentially Theorem XI of [1], which we prove again in slightly altered form in the corollary to Theorem II below.

[13] See [6], pp. 41 and 86.

caused by the existence of monadic extensions can be circumvented, and a completely analogous theorem can be proved.[14] The corollary to Theorem II of this paper (or Theorem XI of [1]) is required for the proof.

4. Preliminary lemmas. As stated above we shall see that all difference fields which have finite incompatible extensions, or a finite monadic extension, have incompatible or monadic extensions of the types shown in the examples; i. e., extensions formed by adjoining solutions of a zero order difference polynomial in one indeterminate. For both incompatible and monadic extensions the proof depends essentially on Lemma III of this section, which in turn depends on Lemma II. It is convenient before considering these lemmas to make some conventions regarding notations.

Let a set of indeterminates $y_1, \cdots, y_n$ be adjoined to a difference field $\mathfrak{F}$. The least difference ring containing $\mathfrak{F}$ and these indeterminates will be denoted by $\mathfrak{F}\{y_1, \cdots, y_n\}$. We denote by y_{ij} the j-th transform of y_i.

Let S be a set of difference polynomials of $\mathfrak{F}\{y_1, \cdots, y_n\}$. The least difference ideal containing the polynomials of S will be denoted by $[S]$, while the perfect difference ideal which they generate is indicated by $\{S\}$.

Now let $\mathfrak{F}$ be any field, not necessarily a difference field. Let indeterminates $y_1, \cdots, y_n$ be adjoined to $\mathfrak{F}$ to form a ring, which does not, of course, contain transforms of the y_i. We denote this ring by $\mathfrak{F}[y_1, \cdots, y_n]$. Let S represent a set of polynomials of this ring. The least ideal containing the polynomials of S will be denoted by $(S)_0$. For definiteness, ideals in $\mathfrak{F}[y_1, \cdots, y_n]$ will be called polynomial ideals. The set of all polynomials of which a power is in $(S)_0$ is the perfect polynomial ideal $\{S\}_0$. It may happen that $\mathfrak{F}$ is a difference field, and the indeterminates of this paragraph are a finite subset of transforms of the indeterminates of a difference ring. Then $(S)_0$ involves only these transforms.

We shall sometimes want to indicate that one of the ideals just described is to be formed from S in an extension $\mathfrak{G}$ of $\mathfrak{F}$. In such a case the symbol $\mathfrak{G}$ is placed in the brackets after S, and separated from it by a semicolon. We write, for example, $\{S; \mathfrak{G}\}$ for the perfect difference ideal generated by polynomials of S in the field $\mathfrak{G}$.

From this point on we depend heavily on the results of [1], [3], [4], and [6]. We remind the reader of the changes in terminology listed in footnote 7.

---

[14] The proof is given in our paper "On extensions of difference fields and the resolvents of prime difference ideals," to be published in the *Proceedings of the American Mathematical Society.*

Our first lemma is a restatement of a well-known theorem concerning algebraic, not difference polynomials.

LEMMA I. *Let* Π *be a prime polynomial ideal with coefficients in a field* $\mathfrak{F}$. *Let* $\mathfrak{G}$ *be an extension of* $\mathfrak{F}$. *Then* $\{\Pi; \mathfrak{G}\}_0$ *is the intersection of prime ideals* $\Pi_1, \Pi_2, \cdots, \Pi_r$, *which are such that any set of parametric indeterminates of* Π *is a set of parametric indeterminates of each* Π_k.

Proof. Let Π involve the indeterminates $u_1, u_2, \cdots, u_q; y_1, y_2, \cdots, y_p$, the u_i forming a set of parametric indeterminates. We form a resolvent [15] for Π using $w = \sum_{i=1}^{p} \mu_i y_i$. The μ_i may be chosen as integers. Let Ω be the prime polynomial ideal $\{\Pi, w - \sum_{i=1}^{p} \mu_i y_i\}_0$, and let

$$\text{L)} \qquad A; A_1, \cdots, A_p$$

be a characteristic set for Ω, with A introducing w, and A_i introducing y_i, $i = 1, \cdots, p$. Each A_i is then of the first degree in y_i.

Let the irreducible factors of A in the field $\mathfrak{G}$ be $B_1, \cdots, B_r$. Each B_k effectively involves w, for otherwise A would not be irreducible in $\mathfrak{F}$.

We adjoin to $\mathfrak{G}$ a set of elements α_i, $i = 1, \cdots, q$, which annul no non-zero polynomial with coefficients in $\mathfrak{G}$.

B_k has a solution $u_i = \alpha_i$, $i = 1, \cdots, q; w = \beta_k$. This solution can annul no initial of any A_j; for these initials are not multiples of B, and hence some linear combination of an initial with B is a non-zero polynomial in the u_i alone. In each A_j we replace the u_i by the α_i, and w by β_k, forming a polynomial $A_j^{(k)}$ of first degree in y_j. By solving $A_j^{(k)}$ we find the value $\gamma_j^{(k)}$ for y_j.

Let Ω_k be the prime polynomial ideal consisting of those polynomials of $\mathfrak{G}[u_1, \cdots, u_q; w; y_1, \cdots, y_p]$ which vanish when we put $u_i = \alpha_i$, $i = 1, \cdots, q$; $w = \beta_k$; $y_j = \gamma_j^{(k)}$, $j = 1, \cdots, p$. Evidently Ω_k contains no polynomial in the u_i alone. We claim that $\{\Omega; \mathfrak{G}\}_0$ is the intersection of the Ω_k.

We see that the solution of Ω_k given in the preceding paragraph annuls the polynomials of the characteristic set L), but not their initials. Hence $\Omega \subset \Omega_k$, and therefore $\{\Omega; \mathfrak{G}\}_0 \subset \Omega_k$.

Let P be a polynomial which is in every Ω_k. We form the remainder R of P with respect to the chain $A_1, \cdots, A_p$. R is a polynomial in the u_i and w alone, which vanishes for the solution $u_i = \alpha_i$, $i = 1, \cdots, q; w = \beta_k$

[15] See [6], p. 86.

of Ω_k. Hence R is a multiple of B_k. Since this is true for each k from 1 to r, R is a multiple of A.

It follows that P has zero remainder with respect to L). We may write P as a finite sum $\sum \delta_i E_i$, where the δ_i are elements of $\mathfrak{G}$ which are linearly independent with respect to $\mathfrak{F}$, and the E_i are polynomials with coefficients in $\mathfrak{F}$. It is easy to see that each E_i has zero remainder with respect to L). Hence the E_i are in Ω, and therefore P is in $\{\Omega; \mathfrak{G}\}_0$. This proves our statement that $\{\Omega; \mathfrak{G}\}_0$ is the intersection of the Ω_k.

Now $\{\Pi; \mathfrak{G}\}_0$ consists of those polynomials of $\{\Omega; \mathfrak{G}\}_0$ which are free of w. For we may write any polynomial with coefficients in $\mathfrak{G}$ in the form $\sum \delta_i E_i$, where the δ_i are elements of $\mathfrak{G}$ which are linearly independent with respect to $\mathfrak{F}$, and the E_i are polynomials with coefficients in $\mathfrak{F}$. The set of all polynomials which have a representation of this form in which every E_i is in Π forms an ideal, and it can easily be seen that this ideal is $(\Pi; \mathfrak{G})_0$. Those polynomials having representations of this form in which every E_i is in Ω form the ideal $(\Omega; \mathfrak{G})_0$. Now Π consists of those polynomials of Ω which are free of w. A polynomial of $(\Omega; \mathfrak{G})_0$ is free of w if and only if each E_i is free of w. Hence $(\Pi; \mathfrak{G})_0$ consists of those polynomials of $(\Omega; \mathfrak{G})_0$ which are free of w. Finally $\{\Pi; \mathfrak{G}\}_0$ consists of polynomials of which a power is in $(\Pi; \mathfrak{G})_0$; and $\{\Omega; \mathfrak{G}\}_0$ consists of polynomials of which a power is in $(\Omega; \mathfrak{G})_0$. Our statement follows immediately from this.

Let Π_k be the prime polynomial ideal consisting of those polynomials of Ω_k which are free of w. Then it follows from the preceding statements that $\{\Pi; \mathfrak{G}\}_0$ is the intersection of the Π_k. The Π_k are prime ideals, and the u_i form a set of parametric indeterminates [16] for each Π_k. This proves the lemma.

LEMMA II. *With the notation of Lemma* I, *if every element of* $\mathfrak{G}$ *not contained in* $\mathfrak{F}$ *is transcendental over* $\mathfrak{F}$, *then* $\{\Pi; \mathfrak{G}\}_0$ *is itself a prime polynomial ideal, and a characteristic set of* Π *is a characteristic set of* $\{\Pi; \mathfrak{G}\}_0$. In particular, if $\mathfrak{F}$ is algebraically closed and $\mathfrak{G}$ is *any* extension of $\mathfrak{F}$, then $\{\Pi; \mathfrak{G}\}_0$ is prime, and a characteristic set of Π is a characteristic set of $\{\Pi; \mathfrak{G}\}_0$.

Proof. To show that $\{\Pi; \mathfrak{G}\}_0$ is prime, it is sufficient to show that A is irreducible in $\mathfrak{G}$; for then $k = 1$ in the proof of Lemma I. Now a poly-

[16] This is so because Ω_k contains no non-zero polynomial in the u_i alone. Hence Π_k contains no non-zero polynomial in the u_i alone. On the other hand Π, and hence each Π_k, contains a non-zero polynomial in $u_1, \cdots, u_q$ and y_j for each j from 1 to p.

nomial S in one indeterminate has the same factors in $\mathfrak{G}$ and in $\mathfrak{F}$, if we consider two factors to be identical when their quotient is an element of the coefficient field. For first of all, it is possible to restrict S and its factors to polynomials with initial coefficient 1. Secondly, S has a factorization into linear factors with initial coefficient 1 in some algebraic extension of $\mathfrak{F}$. All factors of S in any field which meet our restriction are products of these linear factors, and therefore have coefficients algebraic over $\mathfrak{F}$. Hence if such factors have coefficients in $\mathfrak{G}$, then they have coefficients in $\mathfrak{F}$.

By a device due to Kronecker[17] the factors of A can be formed from those of a related polynomial S in one indeterminate. It is easy to see from Kronecker's method of obtaining the factors of A from those of S that, since S has the same factors in $\mathfrak{G}$ as in $\mathfrak{F}$, the same is true of A. Hence A is irreducible in $\mathfrak{G}$, and $\{\Pi; \mathfrak{G}\}_0$ is a prime polynomial ideal.

Let $B_1, \cdots, B_p$ be a characteristic set of Π, and $C_1, \cdots, C_p$ a characteristic set of $\{\Pi; \mathfrak{G}\}_0$. Suppose some C_i is reduced with respect to $B_1, \cdots, B_p$. We let k be the least integer such that C_k is reduced with respect to this chain. We may then take $C_i = B_i$, $1 \leqq i < k$.

We adjoin to $\mathfrak{G}$ elements α_i, $i = 1, \cdots, q$, annulling no polynomial with coefficients in $\mathfrak{G}$. Replacing the u_i by α_i, we solve B_1 for y_1. Substituting the α_i for the u_i in B_2, and the solution just obtained for y_1, B_2 becomes a polynomial in y_2 alone. Solving this and substituting in a similar way into B_3 we obtain a polynomial in y_3 alone. Continuing in this way, we find all the y_i. At no point are we unable to solve; for each ideal $(B_1, \cdots, B_j, I_{j+1}; \mathfrak{F})_0$, where I_{j+1} is the initial of B_{j+1}, contains a non-zero polynomial in the u_i alone,[18] so that I_{j+1} cannot be annulled by the solution obtained for $B_1, \cdots, B_j$. In the same way we see that a solution obtained in this manner annuls no separant of any B_j. Hence when we come to solve B_k, we can find as many distinct solutions as its degree in y_k. In particular, then, we may choose a solution of B_k which does not annul C_k. (Since the chain $B_1, \cdots, B_{k-1}$ is a characteristic set of a prime ideal of polynomials with coefficients in $\mathfrak{G}$, the initial of C_k cannot be annulled by the solution obtained for the u_i and y_j, $j < k$.)

The solution we obtain in this way annuls Π. Hence it annuls $\{\Pi; \mathfrak{G}\}_0$. This contradicts the fact that it does not annul C_k. Hence no C_i is reduced with respect to $B_1, \cdots, B_p$. This proves Lemma II.

The discussion of Example 1 shows that there is no possibility of obtaining

---

[17] See Van der Waerden, *Moderne Algebra*, 1st ed., vol. 1, p. 129.

[18] See [6], Chapter IV, § 9.

a complete analogue of Lemma I for difference polynomials. We can, however, obtain a result analogous to Lemma II.

LEMMA III. *Let $\mathfrak{F}$ be a difference field and Π a reflexive prime difference ideal of polynomials with coefficients in $\mathfrak{F}$. Let $\mathfrak{G}$ be an extension of $\mathfrak{F}$, which is such that every element of $\mathfrak{G}$ not contained in $\mathfrak{F}$ is transcendental over $\mathfrak{F}$, considered as an algebraic field. Then $\{\Pi; \mathfrak{G}\}$ is a reflexive prime difference ideal. Any set of parametric indeterminates of Π is a set of parametric indeterminates of $\{\Pi; \mathfrak{G}\}$, and the effective order of Π in the remaining indeterminates equals the effective order of $\{\Pi; \mathfrak{G}\}$ in those indeterminates.*

In particular, if $\mathfrak{F}$ is algebraically closed, and $\mathfrak{G}$ is *any* extension of $\mathfrak{F}$, the statements of the lemma apply.

Proof. Let $u_1, u_2, \cdots, u_q; y_1, y_2, \cdots, y_p$ be the indeterminates of Π, the u_i forming a set of parametric indeterminates.

The set of polynomials of order not exceeding k in each u_i and y_j will be denoted by Π_k, $k = 0, 1, \cdots$. Π_k is not empty for k sufficiently large, and it is a prime polynomial ideal in the indeterminates u_{im}, $i = 1, \cdots, q$; $m = 0, \cdots, k$, and y_{jn}, $j = 1, \cdots, p$; $n = 0, \cdots, k$.

By Lemma II, the polynomial ideal $\Lambda_k = \{\Pi_k; \mathfrak{G}\}_0$ is evidently a prime ideal in the same set of indeterminates as Π_k. The union Λ of the Λ_k is a prime ideal in the infinite set of indeterminates u_{im}, y_{jn}, where now the second subscripts have the range $0, 1, \cdots$. We shall show that Λ is actually a difference ideal in the u_i and y_j.

Let C be in Λ. We write C as $\sum \gamma_i C_i$, where the γ_i are a finite number of elements of $\mathfrak{G}$ linearly independent with respect to $\mathfrak{F}$, and the C_i are difference polynomials with coefficients in $\mathfrak{F}$.

Now C has zero remainder with respect to the characteristic set of some Λ_k. By Lemma II we may use a characteristic set of Π_k as a characteristic set of Λ_k. Then each C_i has zero remainder with respect to this characteristic set. Hence each C_i is in Π_k, it first transform is in Π_{k+1}, and the transform of C is in Λ_{k+1}, and therefore in Λ.

We shall show that the reflexive prime difference ideal $\{\Lambda\}$ is $\{\Pi; \mathfrak{G}\}$. On the one hand, since every Π_k is in Π, every Λ_k is in $\{\Pi; \mathfrak{G}\}$. Hence $\Lambda \subset \{\Pi; \mathfrak{G}\}$, and therefore $\{\Lambda\} \subset \{\Pi; \mathfrak{G}\}$. On the other hand, since $\Pi \subset \Lambda$, $\{\Pi; \mathfrak{G}\} \subset \{\Lambda; \mathfrak{G}\} = \{\Lambda\}$.

No Λ_k contains a non-zero polynomial in the u_{im} alone. Hence the u_i constitute a set of parametric indeterminates for $\{\Lambda\}$. It remains to show that $\{\Lambda\}$ and Π are of equal effective order in the y_i.

Let A_j be that leader of a characteristic set of Π which introduces y_j. By the result of [1], p. 149, we see that Π is held by no non-zero polynomial in y_1 and the u_i alone which is of lower effective order than A_1 in y_1. By Theorem IX of [1] and the result of [1], p. 149, we see that Π is held by no non-zero polynomial in y_1, y_2 and the u_i, of lower effective order than A_1 in y_1, and of lower effective order than A_2 in y_2. Continuing, we have the following result: Let T designate the set of y_{jn} such that, for each j, n takes those values, and only those, which are less than the effective order of A_j in y_j, and let T_g designate the set consisting of the g-th transforms of the y_{jn} in T. Then Π contains no non-zero difference polynomial in the u_{im} and the y_{jn} of any T_g. By [1], Theorem X, the number of y_{jn} in T is the effective order of Π in the y_j. To prove the lemma it suffices to show that $\{\Lambda\}$ contains no non-zero difference polynomial in the y_{jn} of any T_g and u_i.

No Π_k contains a polynomial in u_{im} and the y_{jn} of a T_g. Hence for all sufficiently large values of k, T_g and the u_{im} present in Π_k form part of a set of parametric indeterminates for Π_k. By Lemma II, they form part of a set of parametric indeterminates for Λ_k. Hence Λ contains no non-zero polynomial in any set T_g and u_i. Then $\{\Lambda\}$ contains no non-zero polynomial in any set T_g and u_i. This proves Lemma III.

We shall need one more result of a different sort, analogous to Gourin's Theorem for differential polynomials.[19]

LEMMA IV. *Let Π and Π' be reflexive prime difference ideals such that Π is properly contained in Π'. Then either Π' is of lower dimension than Π, or it is of the same dimension as Π, every parametric set for Π' is a parametric set for Π, and Π' is of lower effective order than Π relative to any parametric set.*

Proof. Since Π contains no non-zero polynomial in a set of parametric indeterminates of Π', only the last statement of the lemma requires any consideration. Let us assume then that $u_1, \cdots, u_q$ is a parametric set for both Π and Π', the remaining indeterminates being $y_1, \cdots, y_p$.

For non-negative integers a, b, c, let Π_{abc} represent the prime polynomial ideal consisting of all difference polynomials in Π which involve no transform of any y_j of order less than a or more than $a + b$, and no transform of any u_i of order more than c. For a, b, and c sufficiently large, the dimension of Π_{abc} is the sum of the effective order of Π in the y_j and the number $(c + 1)q$ of u_{im} with $m \leqq c$.

---

[19] E. Gourin, "On irreducible systems of algebraic differential equations," *Bulletin of the American Mathematical Society*, vol. 39 (1933), pp. 593-595.

Let polynomial ideals Π'_{abc} be defined for Π' as the Π_{abc} were defined for Π. For all sufficiently large a, b, and c, the dimension of Π'_{abc} is the sum of the effective order of Π' and $(c+1)q$. Now Π'_{abc} contains Π_{abc}. Π' contains a polynomial P which is not in Π. For all sufficiently large a, b, and c, Π'_{abc} contains a transform of P, and hence contains Π_{abc} as a proper subset. Then Π'_{abc} is of lower dimension than Π_{abc} when a, b, and c are large. The lemma follows from the relation above between effective order and dimension.

COROLLARY. *Under the hypotheses of Lemma* III, $\{\Pi; \mathfrak{G}\}$ *contains no polynomial with coefficients in* $\mathfrak{F}$ *except the polynomials of* Π.

Proof. The polynomials of $\{\Pi; \mathfrak{G}\}$ with coefficients in $\mathfrak{F}$ constitute a reflexive prime difference ideal Λ. Since $\Pi \subset \Lambda$, evidently $\{\Pi; \mathfrak{G}\} = \{\Lambda; \mathfrak{G}\}$. By Lemma III, a set of parametric indeterminates for Π is a set of parametric indeterminates for $\{\Pi; \mathfrak{G}\} = \{\Lambda; \mathfrak{G}\}$, and hence is part of a set for Λ; by another application of Lemma III, it is a set for Λ. The effective order of Π relative to this parametric set equals that of Λ by the same lemma. Then, by Lema IV, Λ cannot contain Π as a proper subset. Hence Π and Λ are identical.

5. Incompatible extensions.

THEOREM I. *Let the difference field* $\mathfrak{F}$ *have the incompatible extensions* $\mathfrak{G}$ *and* $\mathfrak{H}$, *of which at least one is a finite extension. Then there is an algebraically irreducible difference polynomial of zero order with coefficients in* $\mathfrak{F}$ *whose general solution is inconsistent.*

Hence, if $\mathfrak{F}$ is algebraically closed it has no incompatible extensions one of which is finite.

Proof. Let $\mathfrak{G}$, say, be a finite extension of $\mathfrak{F}$, and let it arise by the adjunction of elements $\alpha_1, \alpha_2, \cdots, \alpha_r$ to $\mathfrak{F}$. Let Σ be the set of all difference polynomials in unknowns $y_1, y_2, \cdots, y_r$, with coefficients in $\mathfrak{F}$, which are annulled when each y_i is replaced by the corresponding α_i. Σ is not the ideal [0], for otherwise we could find an extension of $\mathfrak{H}$ containing a subfield isomorphic to $\mathfrak{G}$ merely by adjoining to $\mathfrak{H}$ r elements annulling no algebraic difference polynomial with coefficients in $\mathfrak{G}$. Evidently Σ is a reflexive prime difference ideal.

In terms of the ideal Σ we have the following condition for an extension

$\mathfrak{R}$ of $\mathfrak{F}$ to be incompatible with $\mathfrak{G}:\mathfrak{R}$ is incompatible with $\mathfrak{G}$ if and only if $\{\Sigma;\mathfrak{R}\}$ contains a polynomial with coefficients in $\mathfrak{F}$ which is not in Σ.

For suppose, first, that $\mathfrak{R}$ and $\mathfrak{G}$ are not incompatible. Then an extension $\mathfrak{R}'$ of $\mathfrak{R}$ contains a subfield isomorphic to $\mathfrak{G}$. Let $\bar{\alpha}_i$ be the element of $\mathfrak{R}'$ corresponding to α_i. The $\bar{\alpha}_i$ must annul the polynomials of Σ, but no other polynomial with coefficients in $\mathfrak{F}$. Since every solution of Σ in an extension of $\mathfrak{R}$ is a solution of $\{\Sigma;\mathfrak{R}\}$, it follows that $\{\Sigma;\mathfrak{R}\}$ contains no polynomials with coefficients in $\mathfrak{F}$ except the polynomials of Σ.

Next suppose that $\{\Sigma;\mathfrak{R}\}$ contains no polynomial with coefficients in $\mathfrak{F}$ except the polynomials of Σ. Let $\Pi_1, \Pi_2, \cdots, \Pi_r$ be the reflexive prime difference ideals whose intersection is $\{\Sigma;\mathfrak{R}\}$. At least one Π_i contains no polynomial with coefficients in $\mathfrak{F}$, except the polynomials of Σ. For otherwise there would exist for each i, $i = 1, \cdots, r$, a polynomial B_i of Π_i, with coefficients in $\mathfrak{F}$, which is not in Σ. The product of the B_i would be a polynomial with coefficients in $\mathfrak{F}$ which is in $\{\Sigma;\mathfrak{R}\}$ but not in Σ. This would contradict our assumption. Let Π_1, say, contain no polynomials with coefficients in $\mathfrak{F}$ except the polynomials of Σ. We adjoin a generic zero of Π_1 to $\mathfrak{R}$ forming $\mathfrak{R}'$. The subfield of $\mathfrak{R}'$ formed by adjoining the generic zero of Π_1 to $\mathfrak{F}$ is isomorphic to $\mathfrak{G}$. Hence $\mathfrak{R}$ and $\mathfrak{G}$ are not incompatible. This proves our statement of the condition for incompatibility.

Returning to the consideration of the extension $\mathfrak{H}$, we let $\mathfrak{H}^*$ consist of those elements of $\mathfrak{H}$ which are algebraic over $\mathfrak{F}$. We claim that $\mathfrak{H}^*$ and $\mathfrak{G}$ are incompatible.

Suppose that $\mathfrak{H}^*$ and $\mathfrak{G}$ are not incompatible. Then, by the condition just given, $\{\Sigma;\mathfrak{H}^*\}$ contains no polynomials with coefficients in $\mathfrak{F}$ except the polynomials of Σ. It follows as before that, among the reflexive prime difference ideals $\Pi_1, \cdots, \Pi_r$ whose intersection is $\{\Sigma;\mathfrak{H}^*\}$, at least one, say Π_1, contains no polynomial with coefficients in $\mathfrak{F}$ except the polynomials of Σ.

Let Λ denote $\{\Pi_1;\mathfrak{H}\}$. The corollary to Lemma IV shows that Λ contains no polynomial with coefficients in $\mathfrak{H}^*$, except the polynomials of Π_1. Hence Λ contains no polynomial with coefficients in $\mathfrak{F}$, except the polynomials of Σ. Since $\{\Sigma;\mathfrak{H}\} \subset \Lambda$, our criterion for incompatibility shows that $\mathfrak{G}$ and $\mathfrak{H}$ are not incompatible. This contradiction shows that $\mathfrak{G}$ and $\mathfrak{H}^*$ are incompatible.

$\mathfrak{H}^*$ may be an infinite extension of $\mathfrak{F}$, but $\mathfrak{G}$ must be incompatible with a finite zero-order extension [20] of $\mathfrak{F}$. For $\{\Sigma;\mathfrak{H}^*\}$ contains, by the above

[20] A finite zero-order extension of $\mathfrak{F}$ is an extension of the form $\mathfrak{F}\langle\gamma_1, \cdots, \gamma_n\rangle$,

criterion, a polynomial B with coefficients in $\mathfrak{F}$ which is not in $\mathfrak{Z}$. As shown in [4], p. 449, B can be obtained from a finite number of polynomials of $\mathfrak{Z}$ by a finite number of shufflings.[21] These operations involve only a finite number of elements of $\mathfrak{H}^*$. Let $\mathfrak{H}'$ be the difference field formed by adjoining these elements to $\mathfrak{F}$. Then $\{\mathfrak{Z}; \mathfrak{H}'\}$ contains B. Hence $\mathfrak{G}$ and $\mathfrak{H}'$ are incompatible.

$\mathfrak{H}'$ is a finite zero order extension of $\mathfrak{F}$. But this implies that it is a simple extension. For let $c_1, \cdots, c_s$ be the generators of $\mathfrak{H}'$. There is an element c, a linear combination with integer coefficients of the c_i, such that the algebraic field $\mathfrak{F}(c)$ contains the c_i. Then the difference field $\mathfrak{F}\langle c\rangle$ contains the c_i, and hence is $\mathfrak{H}'$.

Since $\mathfrak{H}'$ is a finite extension incompatible with $\mathfrak{G}$, a repetition of the preceding argument shows that $\mathfrak{G}$ contains an element d algebraic over $\mathfrak{F}$ such that $\mathfrak{G}' = \mathfrak{F}\langle d\rangle$ is incompatible with $\mathfrak{H}'$.

We now know that the existence of the incompatible extensions $\mathfrak{G}$ and $\mathfrak{H}$ implies the existence of simple extensions algebraic over $\mathfrak{F}$ which are incompatible. It only remains to show that two such simple extensions can be produced by adjoining different solutions of one irreducible zero-order difference polynomial.

Let P and Q be the algebraically irreducible zero-order difference polynomials annulled by c and d respectively. Considering $\mathfrak{F}$ as an algebraic field, we adjoin to it a solution γ of P and a solution δ of Q, forming the algebraic field $\mathfrak{F}_0$. There is an integer k such that $\mathfrak{F}(\gamma + k\delta)$ is $\mathfrak{F}_0$. Let R be the algebraically irreducible polynomial with coefficients in $\mathfrak{F}$ which is annulled by $\gamma + k\delta$.

Using Theorem IV of [1] in place of the fundamental theorem of algebra, we adjoin to $\mathfrak{H}'$ solutions of the difference polynomial Q equal in number to its degree. Among these solutions there must be a $\bar{d}$ such that the mapping $c \to \gamma$, $\bar{d} \to \delta$ carries $\mathfrak{F}(c, \bar{d})$ isomorphically into $\mathfrak{F}_0$. We denote by $\mathfrak{Z}_1$ the reflexive prime difference ideal consisting of all difference polynomials in the unknown y, and with coefficients in $\mathfrak{F}$, which are annulled by $c + k\bar{d}$. Now $\mathfrak{F}(c + k\bar{d})$ must contain $\mathfrak{F}(c)$. Hence $\mathfrak{F}\langle c + k\bar{d}\rangle$ contains $\mathfrak{F}\langle c\rangle$. In other words, the field $\mathfrak{F}_1$ obtained by adjoining to $\mathfrak{F}$ a generic zero of $\mathfrak{Z}_1$ contains a subfield isomorphic to $\mathfrak{H}'$.

---

where each γ_i is the solution of a zero-order difference polynomial with coefficients in $\mathfrak{F}$. It must not be confused with a finite algebraic extension of $\mathfrak{F}$ considered as an algebraic field, that is, with an extension of the form $\mathfrak{F}(\delta_1, \cdots, \delta_r)$, where the δ_i satisfy algebraic equations with coefficients in $\mathfrak{F}$.

[21] A shuffling, as defined in [4], involves taking transforms, forming linear combinations, and factoring difference polynomials into products of transforms.

The difference polynomial P has a solution $\bar{c}$ in an extension of $\mathfrak{F}\langle d\rangle$, which is such that the mapping $\bar{c} \to \gamma$, $d \to \delta$ carries $\mathfrak{F}(\bar{c}, d)$ isomorphically into $\mathfrak{F}_0$. Let $\mathfrak{Z}_2$ be the reflexive prime difference ideal consisting of all difference polynomials in the unknown y, and with coefficients in $\mathfrak{F}$, which are annulled by $\bar{c} + kd$. We see as above that the field $\mathfrak{F}_2$ obtained by adjoining to $\mathfrak{F}$ a generic zero of $\mathfrak{Z}_2$ contains a subfield isomorphic to $\mathfrak{G}'$. Hence $\mathfrak{F}_1$ and $\mathfrak{F}_2$ are incompatible. Since $\mathfrak{Z}_1$ and $\mathfrak{Z}_2$ are essential prime ideals in the decomposition of $\{R\}$, Theorem I is proved.

COROLLARY. *If the difference field $\mathfrak{F}$ is algebraically closed, there is an extension $\mathfrak{K}$ of $\mathfrak{F}$ such that every reflexive prime difference ideal with coefficients in $\mathfrak{F}$ has a generic zero in $\mathfrak{K}$.*

We suppose that the reflexive prime difference ideals with coefficients in $\mathfrak{F}$ have been well-ordered. The proof is then easily carried out by induction. We produce a correspondence between the ideals and certain extensions of $\mathfrak{F}$, such that the extension corresponding to a reflexive prime difference ideal $\mathfrak{Z}$ contains a generic zero of $\mathfrak{Z}$, and such that if $\mathfrak{Z}_1$ precedes $\mathfrak{Z}_2$ in the well-ordering, then the extension corresponding to $\mathfrak{Z}_2$ includes that corresponding to $\mathfrak{Z}_1$. Evidently the extension corresponding to the prime ideal which is first in the well-ordering can be constructed. Suppose we have constructed such a set of extensions for all the reflexive prime difference ideals preceding an ideal Λ. Then the union $\mathfrak{G}$ of these extensions is itself an extension of $\mathfrak{F}$. By Theorem I, there is an extension $\mathfrak{H}$ of $\mathfrak{G}$ which contains a generic zero of Λ. We let $\mathfrak{H}$ be the extension of $\mathfrak{F}$ corresponding to Λ. The union of all the extensions of the set is the extension $\mathfrak{K}$ of the corollary.

6. Monadic extensions.

THEOREM II. *If the difference field $\mathfrak{F}$ has a finite monadic extension, then it has a monadic extension $\mathfrak{F}\langle\alpha\rangle$, where α is a solution of a zero-order difference polynomial with coefficients in $\mathfrak{F}$.*

Hence, if $\mathfrak{F}$ is algebraically closed it does not have a finite monadic extension.

Proof. Let $\mathfrak{G}$ be a monadic extension of $\mathfrak{F}$ formed by adjoining to $\mathfrak{F}$ the elements α_i, $i = 1, \cdots, n$. The α_i annul some non-zero difference polynomial with coefficients in $\mathfrak{F}$; for otherwise, if we adjoined to $\mathfrak{G}$ a set of elements β_i, $i = 1, \cdots, n$ annulling no non-zero difference polynomial with

coefficients in $\mathfrak{G}$, the resulting extension of $\mathfrak{G}$ would contain an extension of $\mathfrak{F}$ isomorphic to $\mathfrak{G}$ and distinct from $\mathfrak{G}$, namely $\mathfrak{F}\langle\beta_1, \cdots, \beta_n\rangle$.

Let $\mathfrak{G}^*$ be the subfield of $\mathfrak{G}$ consisting of all elements of $\mathfrak{G}$ which are algebraic over the inversive extension [22] $\bar{\mathfrak{F}}$ of $\mathfrak{F}$. We shall show that $\mathfrak{G}$ coincides with $\mathfrak{G}^*$. We assume the contrary and obtain a contradiction.

An element of $\mathfrak{G}$ whose transform is in $\mathfrak{G}^*$ is itself in $\mathfrak{G}^*$. Hence, by our assumption, $\mathfrak{G}$ contains elements no transform of which is in $\mathfrak{G}^*$. Then $\mathfrak{G}$ must be a monadic extension of $\mathfrak{G}^*$ formed by adjoining the α_i. Let Σ be the reflexive prime difference ideal consisting of all polynomials in $y_1, \cdots, y_n$, with coefficients in $\mathfrak{G}^*$, which vanish when the y_i are replaced by the corresponding α_i. Σ is not empty. Since there is an α_i no transform of which is algebraic over $\mathfrak{G}^*$, Σ either has a set of parametric indeterminates which is not empty, or is of effective order different from zero.

Let Λ denote $\{\Sigma; \mathfrak{G}\}$. By Lemma III, Λ is a reflexive prime difference ideal which, like Σ, either has parametric indeterminates, or is not of effective order zero. Hence a generic zero $\beta_1, \cdots, \beta_n$ of Λ is not contained in $\mathfrak{G}$. By the corollary to Lemma IV, Λ contains no polynomial with coefficients in $\mathfrak{G}^*$ except the polynomials of Σ. Hence the extension $\mathfrak{G}\langle\beta_1, \cdots, \beta_n\rangle$ contains a subfield $G^*\langle\beta_1, \cdots, \beta_n\rangle$ isomorphic to and distinct from $\mathfrak{G}$. This contradicts the fact that $\mathfrak{G}$ is a monadic extension of $\mathfrak{G}^*$. Hence $\mathfrak{G} = \mathfrak{G}^*$.

From what we have just shown, every α_i has a transform α_i^* algebraic over $\mathfrak{F}$. $\mathfrak{F}\langle\alpha_1^*, \cdots, \alpha_n^*\rangle$ is a monadic extension of $\mathfrak{F}$. For first of all, there is some α_i, and hence some α_i^*, no transform of which is in $\mathfrak{F}$. Secondly, if there existed an extension $\mathfrak{R}$ of $\mathfrak{F}$ such that $\mathfrak{F}\langle\alpha_1^*, \cdots, \alpha_n^*\rangle$ had two distinct isomorphisms into $\mathfrak{R}$, then $\mathfrak{F}\langle\alpha_1, \cdots, \alpha_n\rangle$ would have two distinct isomorphisms into the inversive extension of $\mathfrak{R}$, which is impossible.

As we showed in the proof of Theorem I, a finite zero-order extension of a difference field is simple. Hence there exists an element α such that $\mathfrak{F}\langle\alpha\rangle = \mathfrak{F}\langle\alpha_1^*, \cdots, \alpha_n^*\rangle$. α is the solution of a zero-order difference polynomial with coefficients in $\mathfrak{F}$. This proves Theorem II.

---

[22] An inversive extension of a difference field is an extension which, if it contains an element h, contains an element g whose transform is h. In particular, we mean by *the* inversive extension $\bar{\mathfrak{F}}$ of a difference field $\mathfrak{F}$, that unique inversive extension all of whose elements have transforms in $\mathfrak{F}$. It existence is shown in [2]. The inversive extension $\bar{\mathfrak{G}}$ of $\mathfrak{G}$ contains $\bar{\mathfrak{F}}$. Those elements of $\bar{\mathfrak{G}}$ which are elements of $\mathfrak{G}$ and which are algebraic over $\bar{\mathfrak{F}}$ are the elements of $\mathfrak{G}$ which we call algebraic over the inversive extension of $\mathfrak{F}$.

COROLLARY. *If $\mathfrak{G}$ is a finite monadic extension of the difference field $\mathfrak{F}$, then every element of $\mathfrak{G}$ is algebraic over the inversive extension of $\mathfrak{F}$.*

Proof. This is equivalent to the statement $\mathfrak{G} = \mathfrak{G}^*$, proved in the preceding demonstration.

An examination of the proof of Theorem II will show that we have made no use of the fact that the monadic extension $\mathfrak{G}$ has no automorphisms onto itself except the identity. We have used only the fact that $\mathfrak{G}$ cannot be isomorphic to two distinct subfields of an extension of $\mathfrak{F}$. Hence, *if $\mathfrak{G}$ is a finite extension of $\mathfrak{F}$ which has this property, then there is a simple zero-order extension of $\mathfrak{F}$ which contains a transform of every element of $\mathfrak{G}$.*

Conversely, if $\mathfrak{G}' = \mathfrak{F}\langle\alpha\rangle$ is a simple zero-order extension of $\mathfrak{F}$, there is an extension $\mathfrak{G}$ of $\mathfrak{G}'$ which has the property just stated. For let P be the algebraically irreducible difference polynomial whose solution is α. P is of zero order in an indeterminate y, and of degree $n \geqq 1$ in y_0. Using Theorem IV of [1] in place of the fundamental theorem of algebra, we may adjoin $n-1$ additional solutions of P to $\mathfrak{G}'$. The resulting field is $\mathfrak{G}$.

This property of zero-order extensions is in no way surprising; we find quite the same situation in algebraic fields and differential fields. What is anomalous is that some of these extensions have no automorphisms except the identity, i. e. are monadic. We have found no other way of characterizing those zero-order extensions which are monadic. At the end of this paper we give a conjecture which, if correct, gives a somewhat narrower class of extensions which includes the finite monadic extensions.

7. Infinite extensions. We obtain conclusions somewhat weaker than those of the preceding sections when we deal with countably infinite extensions of a difference field which are incompatible or monadic. To study these, we must first consider the following situation.

Let $\mathfrak{H}$ be an extension of the difference field $\mathfrak{H}^*$, such that every element of $\mathfrak{H}$ not included in $\mathfrak{H}^*$ is transcendental over $\mathfrak{H}^*$. Let a set of elements $\alpha_1, \cdots, \alpha_n, \alpha_{n+1}$ be adjoined to $\mathfrak{H}$. Let Σ_n denote the reflexive prime difference ideal consisting of all polynomials in $y_1, \cdots, y_n$, with coefficients in $\mathfrak{H}^*$, which vanish when these y_i are replaced by the corresponding α_i. Let Σ_{n+1} consist of all polynomials in $y_1, \cdots, y_n, y_{n+1}$, with coefficients in $\mathfrak{H}^*$, which vanish when these y_i are replaced by the corresponding α_i. Of course Σ_n consist of those polynomials of Σ_{n+1} which are free of y_{n+1}. Let $\Lambda_n = \{\Sigma_n; \mathfrak{H}\}$,

and let $\Lambda_{n+1} = \{\Sigma_{n+1}; \mathfrak{H}\}$. We shall show that Λ_n *consists of those polynomials of* Λ_{n+1} *which are free of* y_{n+1}.

We denote by Σ_{nk} and $\Sigma_{n+1\,k}$, $k = 0, 1, \cdots$, the polynomial ideals consisting of those polynomials of Σ_n and Σ_{n+1}, respectively, which contain no transforms of any y_i of order exceeding k. Let $\Lambda_{nk} = \{\Sigma_{nk}; \mathfrak{H}\}_0$, and let $\Lambda_{n+1\,k} = \{\Sigma_{n+1\,k}; \mathfrak{H}\}_0$. We see from the proof of Lemma III that Λ_n consists of those difference polynomials a transform of which is in some Λ_{nk}, and that Λ_{n+1} consists of those difference polynomials a transform of which is in some $\Lambda_{n+1\,k}$.

Σ_{nk} consists of all polynomials of $\Sigma_{n+1\,k}$ which are free of transforms of y_{n+1}. We choose a complete set T_{nk} of parametric indeterminates for Σ_{nk}. (T_{nk} may be empty, or may consist—when Σ_{nk} is the ideal [0]—of all the indeterminates of Σ_{nk}.) Since $\Sigma_{n+1\,k}$ contains no non-zero polynomial in the indeterminates of T_{nk}, we may select a complete set $T_{n+1\,k}$ of parametric indeterminates for $\Sigma_{n+1\,k}$ which includes T_{nk}. T_{nk} and $T_{n+1\,k}$ are also sets of parametric indeterminates for Λ_{nk} and $\Lambda_{n+1\,k}$ respectively, by Lemma II.

Let Π be the prime polynomial ideal which consists of those polynomials of $\Lambda_{n+1\,k}$ which are free of y_{n+1} and its transforms. Then $\Pi \supset \Lambda_{nk}$. Since Π contains no polynomial in the indeterminates of T_{nk}, it has at least as large a set of parametric indeterminates as Λ_{nk}. Hence Λ_{nk} cannot be a proper subset of Π. Then $\Pi = \Lambda_{nk}$. In other words Λ_{nk} consists of those polynomials of $\Lambda_{n+1\,k}$ which are free of y_{n+1} and its transforms.

Let P be a difference polynomial in Λ_{n+1} which does not involve y_{n+1}. Then for some k, a transform of P is in $\Lambda_{n+1\,k}$, and therefore in Λ_{nk}. Thus P is in Λ_n, which proves our statement.

Let $\gamma_1, \cdots, \gamma_{n+1}$ be a generic zero of Λ_{n+1}. By the result just proved, $\gamma_1, \cdots, \gamma_n$ is a generic zero of Λ_n. Conversely, let $\alpha_1, \cdots, \alpha_n$ be a generic zero of Λ_n. Since $\mathfrak{H}' = \mathfrak{H}\langle\alpha_1, \cdots, \alpha_n\rangle$ is isomorphic to $\mathfrak{H}\langle\gamma_1, \cdots, \gamma_n\rangle$, we can find an extension of $\mathfrak{H}'$ which is isomorphic to $\mathfrak{H}\langle\gamma_1, \cdots, \gamma_n, \gamma_{n+1}\rangle$. Let α_{n+1} be the element corresponding to γ_{n+1} in this isomorphism. Then $\alpha_1, \cdots, \alpha_n, \alpha_{n+1}$ is a generic zero of Λ_{n+1}. Hence *any generic zero of* Λ_n *can be extended to a generic zero of* Λ_{n+1}. It is in this form that our result will be used in studying countably infinite extensions.

THEOREM III. *Let a difference field* $\mathfrak{F}$ *have the at most countably infinite extensions* $\mathfrak{G}$ *and* $\mathfrak{H}$ *which are incompatible. Then* $\mathfrak{F}$ *has incompatible extensions* $\mathfrak{G}^*$ *and* $\mathfrak{H}^*$, *at most countably infinite, every element of which is algebraic over* $\mathfrak{F}$.

Proof. Let $\mathfrak{G}^*$ and $\mathfrak{H}^*$ consist of those elements of $\mathfrak{G}$ and $\mathfrak{H}$, respectively, which are algebraic over $\mathfrak{F}$. We shall assume that $\mathfrak{G}$ and $\mathfrak{H}^*$ are not incompatible, and show that this implies that $\mathfrak{G}$ and $\mathfrak{H}$ are not incompatible.

By our assumption, there is an extension $\mathfrak{H}'$ of $\mathfrak{H}^*$ which contains a subfield isomorphic to $\mathfrak{G}$. This subfield must be formed by adjoining to $\mathfrak{F}$ a set of elements α_i, $i = 1, 2, \cdots$. There is no loss of generality in assuming that there are an infinite number of α_i, since they need not all be distinct.

For each n, $n = 1, 2, \cdots$, let Σ_n denote the reflexive prime difference ideal of polynomials of $\mathfrak{H}^*\{y_1, \cdots, y_n\}$ which vanish when each y_i is replaced by α_i. Let Λ_n be the reflexive prime difference ideal $\{\Sigma_n; \mathfrak{H}\}$. We shall adjoin a set of elements β_i, $i = 1, 2, \cdots$, to $\mathfrak{H}$ such that, for each $n \geqq 1$, $\beta_1, \cdots, \beta_n$ is a generic zero of Λ_n. If Λ_n is $[0]$, this is to mean that $\beta_1, \cdots, \beta_n$ annul no non-zero difference polynomial with coefficients in $\mathfrak{H}$.

For β_1 we choose a generic zero of Λ_1. Suppose we have adjoined $\beta_1, \cdots, \beta_k$. Then $\beta_1, \cdots, \beta_k$ is a generic zero of Λ_k. Now Σ_k consists of those polynomials of Σ_{k+1} which are free of y_{k+1}. By the result obtained at the beginning of this section, we may adjoin an element to the set $\beta_1, \cdots, \beta_k$ to obtain a generic zero of Λ_{k+1}. Let β_{k+1} be this element. By induction we construct the set β_i, $i = 1, 2, \cdots$.

For any n, $\beta_1, \cdots, \beta_n$ annul no polynomial of $\mathfrak{F}\{y_1, \cdots, y_n\}$, except those annulled by $\alpha_1, \cdots, \alpha_n$. For, by the corollary to Lemma IV, Λ_n contains no polynomial with coefficients in $\mathfrak{H}^*$ except the polynomials of Σ_n. Hence Λ_n contains no polynomial with coefficients in $\mathfrak{F}$, except the polynomials of Σ_n with coefficients in $\mathfrak{F}$. It follows that $\mathfrak{H}\langle\beta_1, \beta_2, \cdots\rangle$ contains a subfield isomorphic to $\mathfrak{G}$, namely $\mathfrak{F}\langle\beta_1, \beta_2, \cdots\rangle$. Then $\mathfrak{G}$ and $\mathfrak{H}$ are not incompatible. Since we have obtained a result which contradicts the hypothesis of Theorem III, it follows that $\mathfrak{G}$ and $\mathfrak{H}^*$ are incompatible.

We can find a countable set of elements $\gamma_1, \gamma_2, \cdots$, such that the field $\mathfrak{F}(\gamma_1, \gamma_2, \cdots)$, formed by adjoining the γ_i to $\mathfrak{F}$ considered as an algebraic field, is $\mathfrak{H}$. We need merely take for the γ_i all transforms of the elements which produce $\mathfrak{H}$ when they are adjoined to $\mathfrak{F}$, considered as a difference field. It follows that when $\mathfrak{H}$ is considered as a linear vector space with $\mathfrak{F}$ for coefficient field, its dimension is at most $\aleph_0$. Hence $\mathfrak{H}^*$ is a linear vector space over $\mathfrak{F}$ of dimension at most $\aleph_0$. Therefore $\mathfrak{H}^*$ is an at most countably infinite extension of $\mathfrak{F}$, considered as a difference field. By a repetition of the demonstration above, we see that $\mathfrak{H}^*$ is incompatible with $\mathfrak{G}^*$. This proves Theorem III.

THEOREM IV. *Let a difference field $\mathfrak{F}$ have a monadic extension $\mathfrak{G}$ which is an at most countably infinite extension of $\mathfrak{F}$. Then every element of $\mathfrak{G}$ has a transform which is algebraic over $\mathfrak{F}$.*

Proof. Let $\mathfrak{G}^*$ be the set of elements of $\mathfrak{G}$ which have transforms algebraic over $\mathfrak{F}$. We shall assume that $\mathfrak{G}$ contains elements which are not in $\mathfrak{G}^*$ and obtain a contradiction. $\mathfrak{G}$ can be obtained by adjoining a countable set of elements $\alpha_1, \alpha_2, \cdots$, to $\mathfrak{F}$. Of these elements, at least one is not in $\mathfrak{G}^*$ because of our assumption. Changing the subscripts if necessary, we arrange matters so that α_1 is not in $\mathfrak{G}^*$. Then α_1 annuls no difference polynomial of effective order zero and with coefficients in $\mathfrak{G}^*$. For otherwise, a transform of α_1 would be algebraic over $\mathfrak{F}$, and, α_1 would be in $\mathfrak{G}^*$.

Let Σ_n denote the reflexive prime difference ideal consisting of those polynomials of $\mathfrak{G}^*\{y_1, \cdots, y_n\}$ which vanish when the y_i are replaced by the α_i. Let Λ_n be the reflexive prime difference ideal $\{\Sigma_n; \mathfrak{G}\}$.

We shall adjoin a set of elements $\beta_1, \beta_2, \cdots$ to $\mathfrak{G}$ such that, for each $n \geqq 1$, $\beta_1, \cdots, \beta_n$ is a generic zero of Λ_n. For β_1 we choose a generic zero of Λ_1. Suppose we have already adjoined $\beta_1, \cdots, \beta_k$. Then $\beta_1, \cdots, \beta_k$ is a generic zero of Λ_k. By the result obtained at the begnning of this section, we may adjoin an element β_{k+1} to the set $\beta_1, \cdots, \beta_k$ so as to obtain a generic zero of Λ_{k+1}. By induction we construct the set β_i, $i = 1, 2, \cdots$.

By the provision above concerning α_1, Σ_1 is either [0] or of effective order different from zero. Hence Λ_1 is [0] or of effective order different from zero. Therefore its generic zero β_1 is not in the coefficient field $\mathfrak{G}$.

For each n, Λ_n contains no difference polynomial with coefficients in $\mathfrak{F}$, except the polynomials of Σ_n which have coefficients in $\mathfrak{F}$. Hence $\mathfrak{F}\langle\beta_1, \beta_2, \cdots\rangle$ is isomorphic to $\mathfrak{G}$. Then $\mathfrak{G}\langle\beta_1, \beta_2, \cdots\rangle$ contains two subfields isomorphic to $\mathfrak{G}$, namely $\mathfrak{G}$ itself and $\mathfrak{F}\langle\beta_1, \beta_2, \cdots\rangle$. These subfields are distinct, because β_1 is not in $\mathfrak{G}$. Therefore $\mathfrak{G}$ is not a monadic extension of $\mathfrak{F}$, and we have obtained a contradiction. This proves Theorem IV.[23]

8. Conclusion. We have seen that if a difference field $\mathfrak{F}$ permits finite incompatible extensions, or a finite monadic extension, then there is a single algebraically irreducible difference polynomial A of zero order in its indeter-

[23] As noted after the proof of Theorem II, we have actually proved a theorem concerning a class of extensions wider than the class of monadic extensions. For it is only necessary to assume that $\mathfrak{G}$ is such that no extension of $\mathfrak{F}$ contains two distinct subfields isomorphic to $\mathfrak{G}$.

minate y, such that a monadic extension of $\mathfrak{F}$, or incompatible extensions of $\mathfrak{F}$, can be obtained by adjoining solutions of A.

Now for a single difference polynomial of zero order the two cases of incompatible extensions and of a monadic extension represent opposite extremes of a range of possibilities. Let us consider such a polynomial A, algebraically irreducible, and of degree $n > 1$ in y. We suppose that $\{A\}$ is the intersection of two or more reflexive prime difference ideals $\Pi_1, \cdots, \Pi_r$, whose manifolds we denote by $\mathfrak{M}_1, \cdots, \mathfrak{M}_r$ respectively.

We show as in the proof of Theorem I that there exist extensions of $\mathfrak{F}$ containing n solutions of A. In fact any difference field containing $k < n$ solutions of A can be extended to a field containing n solutions of A. Not all such extensions of $\mathfrak{F}$ need be isomorphic, as we see from Example 1 above. It may be that for a particular polynomial A, the solutions in any one such extension all lie in the same manifold $\mathfrak{M}_i$ (where i depends on the extension). In such a case all the manifolds are evidently incompatible. Examples 1-3 all illustrate this. If, on the other hand, every such extension of $\mathfrak{F}$ contains one solution in each $\mathfrak{M}_i$ (so that we must have $r = n$), then each $\mathfrak{M}_i$ is a monadic manifold. Example 4 illustrates this.

For an example of a distribution of the solutions among the manifolds which is intermediate between these extremes, and which results neither in monadic nor in incompatible extensions, we consider the field $\mathfrak{T}$ which is formed by adjoining a transformally transcendental element u to a difference field $\mathfrak{F}$.

We consider $\mathfrak{T}$ momentarily as an alegbraic field and adjoin a sequence of elements $\alpha_0, \alpha_1, \cdots$, such that $\alpha_i^2 = u_i$, $i = 0, 1, \cdots$. Evidently $\mathfrak{T}(\alpha_0, \alpha_1, \cdots)$ becomes a difference field $(\epsilon)\mathfrak{T}$ if we define the transform of α_i to be $\epsilon_i \alpha_{i+1}$, $i = 0, 1, \cdots$, where each ϵ_i is either $+1$ or -1. All the difference fields $(\epsilon)\mathfrak{T}$ are isomorphic. There is no other way of extending the endomorphism of $\mathfrak{T}$ to $\mathfrak{T}(\alpha_0, \alpha_1, \cdots)$.

Let A be the algebraically irreducible polynomial in y, with coefficients in $\mathfrak{T}$, whose solutions are $\alpha_0 + \alpha_1$, $\alpha_0 - \alpha_1$, $-\alpha_0 + \alpha_1$, $-\alpha_0 - \alpha_1$. We consider A is a difference polynomial. Let γ be any solution of A. Since α_i can be expressed rationally in terms of $\alpha_i + \alpha_{i+1}$ or any of its conjugates, $T\langle\gamma\rangle$ is isomorphic with one of the fields $(\epsilon)\mathfrak{T}$ and therefore contains four solutions of A. Since the fields $(\epsilon)\mathfrak{T}$ are isomorphic, all fields formed by adjoining solutions of A to $\mathfrak{T}$ are isomorphic. Hence the four solutions must be distributed in a fixed manner among the essential irreducible manifolds of A, and there must be at least one of them in each of these manifolds. We

shall show that there are actually two essential irreducible manifolds of A, and that two of the four solutions lie in each of them.

For convenience, let γ be a solution in one of the fields $(\epsilon)\mathfrak{T}$. Then γ_0 is a conjugate of $\alpha_0 + \alpha_1$. We form the algebraic field $\mathfrak{T}_1 = \mathfrak{T}(\gamma_0)$ which must contain α_0 and α_1. α_2 is of second degree over $\mathfrak{T}_1$. Hence γ_1 is of second degree over $\mathfrak{T}_1$. Then the transform A_1 of A factors in $\mathfrak{T}_1$ into two irreducible polynomials of second degree. Let $\mathfrak{T}_h$ denote $\mathfrak{T}(\alpha_0, \alpha_1, \cdots, \alpha_{h-1})$. It follows as with A_1 that the h-th transform A_h of A factors in $\mathfrak{T}_h$ into two algebraically irreducible polynomials of second degree. Theorem IV of [1] now shows that A has two irreducible manifolds $\mathfrak{M}_1$ and $\mathfrak{M}_2$.

From our previous discussion we know that if we adjoin a solution of either $\mathfrak{M}_1$ or $\mathfrak{M}_2$ to $\mathfrak{T}$, we obtain a field isomorphic to any $(\epsilon)\mathfrak{T}$, and containing a solution γ' of $\mathfrak{M}_1$, and a solution γ'' of $\mathfrak{M}_2$. But the mapping which carries each α_i into $-\alpha_i$, or the corresponding mapping in a field isomorphic to $(\epsilon)\mathfrak{T}$, is easily seen to be an automorphism of the difference field $(\epsilon)\mathfrak{T}$. It carries γ' into $-\gamma'$, and γ'' into $-\gamma''$. Hence $-\gamma'$ and $-\gamma''$ are solution of $\mathfrak{M}_1$ and $\mathfrak{M}_2$ respectively. This proves that there are always two solutions in each manifold. Evidently $\mathfrak{M}_1$ and $\mathfrak{M}_2$ are not incompatible, nor does either furnish a monadic extension of $\mathfrak{T}$.

We may duplicate this example in the field $\mathfrak{R}$ of rational functions of x with complex coefficients, the transform of a function $f(x)$ being defined as $f(x+1)$. We simply replace the u_k in the preceding example by $x + k$.

It would be interesting to know whether the fields $\mathfrak{R}$ or $\mathfrak{T}$ have finite monadic or incompatible extensions (other than those obtainable from incompatible or monadic extensions of $\mathfrak{F}$ by adjoining x or u). It would follow that they do not if we could establish the following *conjecture*: If an essential irreducible manifold $\mathfrak{M}$ of an algebraically irreducible zero order difference polynomial A with coefficients in a difference field $\mathfrak{F}$ is monadic, or if it is incompatible with another manifold of A, then a solution of $\mathfrak{M}$ is contained in an extension of $\mathfrak{F}$ which, considered as an algebraic field, is a finite extension. In other terms, we conjecture that the characteristic set of A which is annulled by $\mathfrak{M}$ terminates in a polynomial which is of first degree in the highest transform of the indeterminate present in it.

RUTGERS UNIVERSITY.

BIBLIOGRAPHY.

[1] R. M. Cohn, "Manifolds of difference polynomials," *Transactions of the American Mathematical Society*, vol. 64 (1948), pp. 133-172.

[2] ———, "Inversive difference fields," *Bulletin of the American Mathematical Society*, vol. 55 (1949), pp. 597-603.

[3] J. L. Doob and J. F. Ritt, "Systems of algebraic difference equations," *American Journal of Mathematics*, vol. 55 (1933), pp. 505-514.

[4] H. W. Raudenbush and J. F. Ritt, "Ideal theory and algebraic difference equations," *Transactions of the American Mathematical Society*, vol. 46 (1939), pp. 445-452.

[5] J. F. Ritt, "Complete difference ideals," *American Journal of Mathematics*, vol. 63 (1941), pp. 681-690.

[6] ———, "Differential algebra," *Colloquium Publications of the American Mathematical Society*, vol. 33.